

03/30/00
J6644 U.S. PTO

JC598 U.S. PTO
09/537710
03/30/00

Please type a plus sign (+) inside this box-> + /

UTILITY	Atty Doc. No. <u>3377/99-Util</u>	Total Page <u>59</u>
PATENT APPLICATION	FIRST NAMED INVENTOR OR APPLICATION IDENTIFIER	
TRANSMITTAL	Anders DAHLQVIST	
	Express Mail Label No. _____	

Application Elements

Address To: Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

1. / X / Fee transmittal Form
(Submit an original, and a duplicate for fee processing)
2. / X / Specification Total Pages /
(Preferred arrangement set for below)

6. / / Microfiche Computer Program (Appendix)
/7. / Nucleotide and/or Amino Acid Sequence Submission
(if applicable, all necessary)

Descriptive title of the Invention
Cross References to Related Application
Statement Regarding Fed. Sponsored R & D
Reference to Microfiche Appendix
Background of the Invention
Brief Summary of the Invention
Brief Description of the Drawings (if filed)
Detailed Description
Claim(s)
Abstract of the Disclosure

- a./ / Computer Readable Copy
b/ / Paper Copy (Identical to computer copy)
c/ / Statement verifying identity of above copies
ACCOMPANYING APPLICATIONS PARTS
8./ / Assignment Papers (cover sheet & document(s)
9/ / 37 CFR 3.73(b)Statement / / Power of Attorney
10./ / English Translation Document (if applicable)
11./ / Information Disclosure / / Copies of IDS Citations
12./ / Preliminary Amendment
13./ x/Return Receipt Postcard (MPEP 503)

3. / X / Drawing(s)(35 USC 113)(Figs.) Total Sheets /6 /

Should be specifically itemized)
14./ / Small Entity / / Statement filed in prior application
Statements Status still proper and desired
15./ / Certified Copy of Priority Document(s)
(if foreign priority is claimed)

4./ /Oath or Declaration Total Pages / /
a / / Newly executed (original or copy)
b./ /Copy from a prior application (37 CFR 1.63(d)
(For Continuation/Divisional with Box 17 completed)
Note Box 5 below

i./ / DELETION OF INVENTOR(S)

Signed statement attached deleting
inventor(s) named in the prior application
see 37 CFR 1.63(d)(2) and 1.33(b).

5. / / Incorporation by reference (useable if Box 4b is checked)
The entire disclosure of the prior application, from which a
copy of the oath or declaration is supplied under Box 4b
is considered as being part of the disclosure of the accompanying
application and is hereby incorporated by reference therein.

16./ / Other _____

17. If a Continuing Application, check appropriate box and supply the requisite information:

/ X /Continuation / /Divisional / / Continuation-in part (CIP) of prior application No. 06/180,687

CORRESPONDENCE ADDRESS

/ / Customer Number or Bar code Label

or / / Correspondence address below

Insert Customer No. or Attach bar code label here

Name: Herbert B. Keil
KEIL & WEINKAUF

Address: 1101 Connecticut Ave., N.W.

City Washington State: D.C. Zip Code 20036

Country USA Telephone: (202)659-0100 Fax: (202)659-0105

The filing fee has been calculated as shown below:

For:	Number Filed	Number Extra	SMALL/LARGE ENTITY	BASIC FEE
				\$345./\$690.

Basic Fee..... \$ 690.00

Total Claims: 30 -20 = _____ x \$09./\$18. = 180.00

Indep. Claims: _____ -3 = _____ x \$39./\$78. = _____

[X] Multiple Dependent Claim(s) presented: \$130./260 = 260.00

[x] A check is enclosed for the filing fee. \$ 1,130.00

*If the difference is less than zero, enter "0".

[X] A check for \$1,130.00 for the filing fee and recordation fee.

[X] The Commissioner is hereby authorized to charge any other fee required, including the issue fee, in connection with the filing and prosecution of this application, and to the extent necessary, applicant(s) hereby petition for extension(s) of time under 37 CFR 1.136, to be charged to our Deposit Account 11-0345.

Respectfully submitted,
KEIL & WEINKAUF



Herbert B. Keil
Reg. No. 18,967

1101 Connecticut Ave., N.W.
Washington, D.C. 20036
(202)659-0100

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:)
DAHLQUIST et al.) Group Art:
Serial No. Not yet assigned) Examiner
Filed: With Application)
)

For: A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE
PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA
MOLECULES ENCODING THESE ENZYMES

PRELIMINARY AMENDMENT

Hon. Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

Prior to examination, kindly amend the above-identified application as follows:

IN THE SPECIFICATION

Page 1, after the title, insert

--This is a continuation of provisional application Serial No. 60/180,687, filed February
7, 2000.--

IN THE CLAIMS

Claim 3, line 1, delete "claims 1 or 2" and insert --claim 1--.

Claim 4, line 1, delete "claims 1 to 3 and insert --claim 1--.

Claim 5, line 1, delete "claims 1 to 4" and insert --claim 1--.

Claim 6, line 1, delete "claims 1 to 5" and insert --claim 1--.

Claim 9, line 1, delete "claims 7 or 8" and insert --claim 7--.

Claim 10, line 2, delete "claims 7 to 9" and insert --claim 7--.

Claim 11, line 1, delete "claims 7 to 10" and insert --claim 7--.

Claim 12, lines 1 and 2, delete "claims 7 to 11" and insert --claim 7--

Claim 13, lines 1 and 2, delete "claims 7 to 11 or a gene construct according to claim 12" and insert --claim 7--.

Claim 15, line 1, delete "claims 13 or 14" and insert --claim 13--.

Claim 16, line 2, delete "claims 7 to 11" and insert --claim 7- -;

line 3, delete "to 15" and insert --28--.

Claim 18, line 1, delete "claims 16 or 17" and insert --claim 16--.

Claim 19, line 1, delete "claims 16 to 18" and insert --claim 16--.

Claim 20, line 1, delete "claims 16 to 19" and insert --claim 19--.

Claim 21, line 1, delete "claims 16 to 20" and insert --claim 16--.

Claim 22, line 1, delete "claims 16 to 21" and insert --claim 16--.

Claim 23, line 1, delete "claims 16 to 22" and insert --claim 16--.

Claim 25, line 1, change "24" to ---30--.

Claim 15, line 1, delete "24" and insert --20--.

Cancel claims 24, 26 and 27.

Insert the following new claims.

--28. A vector comprising the gene construct of claim 12.

29. A vector according to claim 28, further comprising a selectable master gene and/or nucleotide sequences for the replication in a host cell or the integration into the genome of the host cell.

30. A process for the production of triacylglycerol, comprising growing a transgenic cell or organism according to claim 16 under conditions whereby a nucleotide sequence encoding an enzyme catalyzing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol

is expressed and whereby said transgenic cells comprising an enzyme catalyzing in a acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.

31. A method of producing triacylglycerol and/or triacylglycerol with uncommon fatty acids which comprises transforming an organism or host cell using the nucleotide sequence of claim 7, whereby the transformation results in an altered, preferably, increased oil content of the cell or organism.

32. A method of producing triacylglycerol and/or triacylglycerols with uncommon fatty acids using the nucleotide sequence of claim 7.

33. A method of producing triacylglycerol and/or triacylglycerols with uncommon fatty acids using the enzyme of claim 1.

R E M A R K S

The claims have been amended to put the application in better form for U.S. filing. No new matter has been added.

Entry of the above amendment is respectfully solicited.

Respectfully submitted,

KEIL & WEINKAUF


Herbert B. Keil
Reg. No. 18,967

1101 Connecticut Ave., N.W.
Washington, D.C. 20036
(202)659-0100

A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE
PRODUCTION OF TRIACYLGLYCEROL AND RECOMBINANT DNA
MOLECULES ENCODING THESE ENZYMES

- 5 The present invention relates to the isolation, identification and characterization of recombinant DNA molecules encoding enzymes catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.
- 10 Triacylglycerol (TAG) is the most common lipid-based energy reserve in nature. The main pathway for synthesis of TAG is believed to involve three sequential acyl-transfers from acyl-CoA to a glycerol backbone (1, 2). For many years, acyl-CoA : diacylglycerol acyltransferase (DAGAT), which catalyzes the third acyl transfer reaction, was thought to be the only unique enzyme involved in
- 15 TAG synthesis. It acts by diverting diacylglycerol (DAG) from membrane lipid synthesis into TAG (2). Genes encoding this enzyme were recently identified both in the mouse (3) and in plants (4, 5), and the encoded proteins were shown to be homologous to acyl-CoA : cholesterol acyltransferase (ACAT). It was also recently reported that another DAGAT exists in the oleaginous fungus
- 20 *Mortierella ramanniana*, which is unrelated to the mouse DAGAT, the ACAT gene family or to any other known gene (6).

The instant invention relates to novel type of enzymes and their encoding genes for transformation. More specifically, the invention relates to use of a

25 type of genes encoding a not previously described type of enzymes hereinafter designated phospholipid:diacylglycerol acyltransferases (PDAT), whereby this enzyme catalyses an acyl-CoA-independent reaction. The said type of genes expressed alone in transgenic organisms will enhance the total amount of oil (triacylglycerols) produced in the cells. The PDAT genes, in combination with a

30 gene for the synthesis of an uncommon fatty acid will, when expressed in transgenic organisms, enhance the levels of the uncommon fatty acids in the triacylglycerols.

There is considerable interest world-wide in producing chemical feedstock, such as fatty acids, for industrial use from renewable plant resources rather than non-renewable petrochemicals. This concept has broad appeal to
5 manufacturers and consumers on the basis of resource conservation and provides significant opportunity to develop new industrial crops for agriculture.

There is a diverse array of unusual fatty acids in oils from wild plant species and these have been well characterised. Many of these acids have industrial
10 potential and this has led to interest in domesticating relevant plant species to enable agricultural production of particular fatty acids.

Development in genetic engineering technologies combined with greater understanding of the biosynthesis of unusual fatty acids now makes it possible
15 to transfer genes coding for key enzymes involved in the synthesis of a particular fatty acid from a wild species into domesticated oilseed crops. In this way individual fatty acids can be produced in high purity and quantities at moderate costs.
20 In all crops like rape, sunflower, oilpalm etc., the oil (i.e. triacylglycerols) is the most valuable product of the seeds or fruits and other compounds like starch, protein, and fibre is regarded as by-products with less value. Enhancing the quantity of oil per weight basis at the expense of other compounds in oil crops would therefore increase the value of crop. If genes, regulating the allocation of
25 reduced carbon into the production of oil can be up-regulated, the cells will accumulate more oil on the expense af other products. Such genes might not only be used in already high oil producing cells such as oil crops but could also induce significant oil production in moderate or low oil containing crops such as e.g. soy, oat, maize, potato, sugarbeets, and turnips as well as in micro-
30 organisms.

Summary of the invention

Many of the unusual fatty acids of interest, e.g. medium chain fatty acids, hydroxy fatty acids, epoxy fatty acids and acetylenic fatty acids, have physical properties that are distinctly different from the common plant fatty acids. The present inventors have found that, in plant species naturally accumulating these uncommon fatty acids in their seed oil (i.e. triacylglycerol), these acids are absent, or present in very low amounts in the membrane (phospho)lipids of the seed. The low concentration of these acids in the membrane lipids is most likely a prerequisite for proper membrane function and thereby for proper cell functions. One aspect of the invention is that seeds of transgenic crops can be made to accumulate high amounts of uncommon fatty acids if these fatty acids are efficiently removed from the membrane lipids and channelled into seed triacylglycerols.

The inventors have identified a novel class of enzymes in plants catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the production of triacylglycerol through an acyl-CoA-independent reaction and that these enzymes (phospholipid:diacylglycerol acyltransferases abbreviated as PDAT) are involved in the removal of hydroxylated, epoxigenated fatty acids, and probably also other uncommon fatty acids such as medium chain fatty acids, from phospholipids in plants.

This enzyme reaction was shown to be present in microsomal preparations from baker's yeast (*Saccharomyces cerevisiae*). The instant invention further pertains to an enzyme comprising an amino acid sequence as set forth in SEQ ID No. 2 or a functional fragment, derivate, allele, homologue or isoenzyme thereof. A so called 'knock out' yeast mutant, disrupted in the respective gene was obtained and microsomal membranes from the mutant was shown to totally lack PDAT activity. Thus, it was proved that the disrupted gene encodes for a PDAT enzyme (SEQ ID NO. 1 and 2).

The instant invention pertains further to an enzyme comprising an amino acid sequence as set forth in SEQ ID NO. 1a, 2b or 5a or a functional fragment, derivate, allele, homologue or isoenzyme thereof.

Further genes and/or proteins of so far unknown function were identified and
5 are contemplated within the scope of the instant invention. A gene from Schizosaccharomyces pombe, SPBC776.14 (SEQ ID. NO. 3), a putative open reading frame CAA22887 of the SPBC776.14 (SEQ ID NO. 13) were identified. Further Arabidopsis thaliana genomic sequences (SEQ ID NO. 4, 10 and 11)
10 coding for putative proteins were identified, as well as a putative open reading frame AAC80628 from the A. thaliana locus AC 004557 (SEQ ID NO. 14) and a putative open reading frame AAD10668 from the A. thaliana locus AC 003027 (SEQ ID NO. 15) were identified.

Also, a partially sequenced cDNA clone from Neurospora crassa (SEQ ID NO. 9) and a Zea mays EST (Extended Sequence Tac) clone (SEQ ID NO. 7) and
15 corresponding putative amino acid sequence (SEQ ID NO. 8) were identified.

Finally, two cDNA clones were identified, one Arabidopsis thaliana EST (SEQ ID NO. 5 and corresponding predicted amino acid sequence SEQ ID NO. 6) and a Lycopersicon esculentum EST clone (SEQ ID NO. 12) were identified. Further, enzymes designated as PDAT comprising an amino acid sequence
20 selected from the group consisting of sequences as set forth in SEQ ID NO 2a, 3a, 5b, 6 or 7b are contemplated within the scope of the invention. Moreover, an enzyme comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or homologue thereof selected from the group consisting of sequences as set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a,
25 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12 or a functional fragment, derivate, allele, homologue or isoenzyme of the enzyme encoding amino acid sequence are included within the scope of the invention.

A functional fragment of the instant enzyme is understood to be any polypeptide
30 sequence which shows specific enzyme activity of a phospholipid:diacylglycerol acyltransferase (PDAT). The length of the functional fragment can for example

vary in a range from about 660 ± 10 amino acids to 660 ± 250 amino acids, preferably from about 660 ± 50 to 660 ± 100 amino acids, whereby the „basic number“ of 660 amino acids corresponds in this case to the polypeptide chain of the PDAT enzyme of SEQ ID NO. 2 encoded by a nucleotide sequence according to SEQ ID NO. 1. Consequently, the „basic number“ of functional fulllength enzyme can vary in correspondance to the encoding nucleotide sequence.

A portion of the instant nucleotide sequence is meant to be any nucleotide sequence encoding a polypeptid which shows specific activity of a phospholipid:diacylglycerol acyltransferase (PDAT). The length of the nucleotide portion can vary in a wide range of about several hundereds of nucleotides based upon the coding region of the gene or a highly conserved sequenunce. For example the length varies in a range form about 1900 ± 10 to 1900 ± 1000 nucleotides, preferably form about 1900 ± 50 to 1900 ± 700 and more preferably form about 1900 ± 100 to 1900 ± 500 nucleotides. whereby the „basic number“ of 1900 nucleotdies corresponds in this case to the encoding nucleotide sequence of the PDAT enzyme of SEQ ID NO. 1. Consequently, the „basic number“ of functional fulllength gene can vary.

An allelic variant of the instant nucleotide sequence is understood to be any different nucleotide sequence which encodes a polypeptide with a functionally equivalent function. The alleles pertain naturally occuring variante of the instant nucleotide sequences as well as synthetic nucleotide sequences produces by methods known in the art. Contemplated are even altered nucleotide sequences which result in an enzyme with altered activity and/or regulation or which is resistant against specific inhibitors. The instant invention further includes natural or synthetic mutations of the originally isolated nucleotide sequences. These mutations can be substitution, addition, deletion, inversion or insertion of one or more nucleotides.

A homologues nucleotide sequence is understood to be a complementary sequence and/or a sequence which specifically hybridizes with the instant nucleotide sequence. Hybridizing sequences include similar sequences selected from the group of DNA or RNA which specifically interact to the instant
5 nucleotide sequences under at least moderate stringency conditions which are known in the art. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-
10 65°C. This further includes short nucleotide sequences of e.g. 10 to 30 nucleotides, preferably 12 to 15 nucleotides. Included are also primer or hybridization probes.

A homologue nucleotide sequence included within the scope of the instant invention is a sequence which is at least about 40%, preferably at least about
15 50 % or 60%, and more preferably at least about 70%, 80% or 90% and most preferably at least about 95%, 96%, 97%, 98% or 99% or more homologous to a nucleotide sequence of SEQ ID NO. 1.

All of the aforementioned definitions are true for amino acid sequences and functional enzymes and can easily transferred by a person skilled in the art.
20

Isoenzymes are understood to be enzymes which have the same or a similar substrate specificity and/or catalytic activity but a different primary structure.

In a first embodiment, this invention is directed to nucleic acid sequences that encode a PDAT. This includes sequences that encode biologically active PDATs as well as sequences that are to be used as probes, vectors for transformation or cloning intermediates. The PDAT encoding sequence may encode a complete or partial sequence depending upon the intended use. All or a portion of the genomic sequence, cDNA sequence, precursor PDAT or mature PDAT is intended.
25
30

Further included is a nucleotide sequence selected from the group consisting of sequences set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 9b, 10, 10b or 11 or a portion, derivate, allele or homologue thereof. The invention pertains a partial 5 nucleotide sequence corresponding to a fulllength nucleotide sequence selected from the group consisting of sequences set forth in SEQ ID No. 5, 5b, 6b, 7, 8b, 9, 11b or 12 or a portion, derivate, allele or homologue thereof. Moreover, a nucleotide sequence comprising a nucleotide sequence which is at least 40% homologous to a nucleotide sequence selected form the group consisting of 10 those sequences set forth in SEQ ID No. 1 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12 is contemplated within the scope of the invention.

The instant invention pertains to a gene construct comprising a said nucleotide sequences of the instant invention which is operably linked to a heterologous 15 nucleic acid.

The term operably linked means a serial organisation e.g. of a promotor, coding sequence, terminator and/or further regulatory elements whereby each element can fulfill its original function during expression of the nucleotide sequence.

20 Further, a vector comprising of a said nucleotide sequence of the instant invention is contemplated in the instant invention. This includes also an expression vector as well as a vector further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell and/or the integration into the genome of the host cell.

25 In a different aspect, this invention relates to a method for producing a PDAT in a host cell or progeny thereof, including genetically engineered oil seeds, yeast and moulds or any other oil accumulating organism, via the expression of a construct in the cell. Cells containing a PDAT as a result of the production of 30 the PDAT encoding sequence are also contemplated within the scope of the invention.

Further, the invention pertains a transgenic cell or organism containing a said nucleotide sequence and/or a said gene construct and/or a said vector. The object of the instant invention is further a transgenic cell or organism which is
5 an eucaryotic cell or organism. Preferably, the transgenic cell or organism is a yeast cell or a plant cell or a plant. The instant invention further pertains said transgenic cell or organism having an altered biosynthetic pathway for the production of triacylglycerol. A transgenic cell or organism having an altered oil content is also contemplated within the scope of this invention.

10 Further, the invention pertains a transgenic cell or organism wherein the activity of PDAT is altered in said cell or organism. This altered activity of PDAT is characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme. Moreover, a transgenic cell or organism is
15 included in the instant invention, wherein the altered biosynthetic pathway for the production of triacylglycerol is characterized by the prevention of accumulation of undesirable fatty acids in the membrane lipids.

20 In a different embodiment, this invention also relates to methods of using a DNA sequence encoding a PDAT for increasing the oil-content within a cell.

Another aspect of the invention relates to the accommodation of high amounts of uncommun fatty acids in the triacylglycerol produced within a cell, by introducing a DNA sequence producing a PDAT that specifically removes these
25 fatty acids from the membrane lipids of the cell and channel them into triacylglycerol. Plant cells having such a modification are also contemplated herein.

30 Further, the invention pertains a process for the production of triacylglycerol, comprising growing a said transgenic cell or organism under conditions whereby the said nucleotide sequence is expressed and whereby the said

transgenic cells comprising an said enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol forming triacylglycerol.

Moreover, triacylglycerols produced by the aforementioned process are
5 included in scope of the instant invention.

Object of the instant invention is further the use of an instant nucleotide sequence and/or a said enzyme for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids. The use of a said instant nucleotide
10 sequence and/or a said enzyme of the instant invention for the transformation of any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism is also contemplated within the scope of the instant invention.

15 A PDAT of this invention includes any sequence of amino acids, such as a protein, polypeptide or peptide fragment obtainable from a microorganism, animal or plant source that demonstrates the ability to catalyse the production of triacylglycerol from a phospholipid and diacylglycerol under enzyme reactive conditions. By „enzyme reactive conditions,“ is meant that any necessary
20 conditions are available in an environment (e.g., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

Other PDATs are obtainable from the specific sequences provided herein. Furthermore, it will be apparent that one can obtain natural and synthetic
25 PDATs, including modified amino acid sequences and starting materials for synthetic-protein modelling from the exemplified PDATs and from PDATs which are obtained through the use of such exemplified sequences. Modified amino acid sequences include sequences that have been mutated, truncated, increased and the like, whether such sequences were partially or wholly
30 synthesised. Sequences that are actually purified from plant preparations or are identical or encode identical proteins thereto, regardless of the method

used to obtain the protein or sequence, are equally considered naturally derived.

Further, the nucleic acid probes (DNA and RNA) of the present invention can
5 be used to screen and recover „homologous“ or „related“ PDATs from a variety of plant and microbial sources.

Further, it is also apparent that a person skilled in the art can, with the information provided in this application, in any organism identify a PDAT
10 activity, purify an enzyme with this activity and thereby identify a „non-homologues“ nucleic acid sequence encoding such an enzyme.

The present invention can be essentially characterized by the following aspects:

- 15 1. Use of a PDAT gene (genomic clone or cDNA) for transformation.
2. Use of a DNA molecule according to item 1 wherein said DNA is used for transformation of any organism in order to be expressed in this organism and result in an active recombinant PDAT enzyme in order to increase oil content of the organism.
- 20 3. Use of a DNA molecule of item 1 wherein said DNA is used for transformation of any organism in order to prevent the accumulation of undesirable fatty acids in the membrane lipids.
4. Use according to item 1, wherein said PDAT gene is used for transforming transgenic oil accumulating organisms engineered to produce any uncommon fatty acid which is harmful if present in high amounts in membrane lipids, such as medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.
- 25 5. Use according to item 1, wherein said PDAT gene is used for transforming organisms, and wherein said organisms are crossed with other oil accummulating organisms engineered to produce any uncommon fatty acid

which is harmful if present in high amounts in membrane lipids, comprising medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.

6. Use according to item 1, wherein the enzyme encoded by said PDAT gene or cDNA is coding for a PDAT with distinct acyl specificity.
- 5 7. Use according to item 1 wherein said PDAT encoding gene or cDNA, is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 30% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 10 8. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 40% or more *identical* to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 15 9. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 60% or more *identical* to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 20 10. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cerevisiae*, or contain nucleotide sequences coding for an amino acid sequence 80% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
- 25 11. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from plants or contain nucleotide sequences coding for an amino acid sequence 40% or more identical to the amino acid sequence of PDAT from *Arabidopsis thaliana* or to the protein encoded by the fulllength counterpart of the partial *Zea mays*, *Lycopericon esculentum*, or *Neurospora crassa* cDNA clones.
- 30 12. Transgenic oil accumulating organisms comprising, in their genome, a PDAT gene transferred by recombinant DNA technology or somatic hybridization.

13. Transgenic oil accumulating organisms according to item 12 comprising, in
their genome, a PDAT gene having specificity for substrates with a
particular uncommon fatty acid and the gene for said uncommon fatty acid.
14. Transgenic organisms according to item 12 or 13 which are selected from
5 the group consisting of fungi, plants and animals.
15. Transgenic organisms according to item 12 or 13 which are selected from
the group of agricultural plants.
16. Transgenic organisms according to item 12 or 13 which are selected from
the group of agricultural plants and where said PDAT gene is expressed
10 under the control of a storage organ specific promotor.
17. Transgenic organisms according to item 12 or 13 which are selected from
the group of agricultural plants and where said PDAT gene is expressed
under the control of a seed promotor.
18. Oils from organisms according to item 12 – 17.
- 15 19. A method for altering acyl specificity of a PDAT by alteration of the
nucleotide sequence of a naturally occurring encoding gene and as a
consequence of this alternation creating a gene encoding for an enzyme
with novel acyl specificity.
20. A protein encoded by a DNA molecule according to item 1 or a functional
fragment thereof.
21. A protein of item 20 designated phospholipid:diacylglycerol acyltransferase.
22. A protein of item 21 which has a distinct acyl specificity.
23. A protein of item 13 having the amino acid sequence as set forth in SEQ,
ID NO. 2, 13, 14 or 15 (and the proteins encoded by the fulllength or partial
25 genes set forth in SEQ. ID. NO. 1, 3, 4, 5, 7, 9, 10, 11 or 12) or an amino
acid sequence with at least 30 % homology to said amino acid sequence.
24. A protein of item 23 isolated from *Saccharomyces cereviseae*.

30 General methods:

Yeast strains and plasmids. The wild type yeast strains used were either FY1679 (*MAT α his3-Δ200 leu2-Δ1 trp1-Δ6 ura3-52*) or W303-1A (*MAT α ADE2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1*) (7). The YNR008w::KanMX2 disruption strain FVKT004-04C(AL), which is congenic to FY1679, was obtained from the Euroscarf collection (8). A 2751 bp fragment containing the YNR008w gene with 583 bp of 5' and 183 bp of 3' flanking DNA was amplified from W303-1A genomic DNA using *Taq* polymerase with 5'-TCTCCATCTTCTGCAAAACCT-3' and 5'-CCTGTCAAAAACCTTCTCCTC-3' as primers. The resulting PCR product was purified by agarose gel electrophoresis and cloned into the *EcoRV* site of pBluescript (pbluescript-pdat). For complementation experiments, the cloned fragment was released from pBluescript by *Hind*III-*Sac*I digestion and then cloned between the *Hind*III and *Sac*I sites of pFL39 (9), thus generating pUS1. For overexpression of the PDAT gene, a 2202 bp *EcoRI* fragment from the pBluscript plasmid which contains only 24 bp of 5' flanking DNA was cloned into the *Bam*HI site of the *GAL1-TPK2* expression vector pJN92 (12), thus generating pUS4.

Microsomal preparations. Microsomes from developing seeds of sunflower (*Helianthus annuus*), *Ricinus communis* and *Crepis palaestina* were prepared using the procedure of Stobart and Stymne (11). To obtain yeast microsomes, 1g of yeast cells (fresh weight) was re-suspended in 8 ml of ice-cold buffer (20 mM Tris-Cl, pH 7.9, 10 mM MgCl₂, 1 mM EDTA, 5 % (v/v) glycerol, 1 mM DTT, 0.3 M ammonium sulfate) in a 12 ml glass tube. To this tube, 4 ml of glass beads (diameter 0.45-0.5 mm) were added, and the tube was then heavily shaken (3 x 60 s) in an MSK cell homogenizer (B. Braun Melsungen AG, Germany). The homogenized suspension was centrifuged at 20,000 x g for 15 min at 6°C and the resulting supernatant was again centrifuged at 100,000 x g for 2 h at 6°C. The 100,000 x g pellet was resuspended in 0.1 M potassium phosphate (pH 7.2), and stored at -80°C. It is subsequently referred to as the crude yeast microsomal fraction.

Lipid substrates. Radio-labeled ricinoleic (12-hydroxy-octadecenoic) and vernolic (12,13-epoxy-octadecenoic) acids were synthesized enzymatically from [1-¹⁴C]oleic acid and [1-¹⁴C]linoleic acid, respectively, by incubation with 5 microsomal preparations from seeds of *Ricinus communis* and *Crepis palaestina*, respectively (12). The synthesis of phosphatidylcholines (PC) or phosphatidylethanolamines (PE) with ¹⁴C-labeled acyl groups in the *sn*-2 position was performed using either enzymatic (13), or synthetic (14) acylation 10 of [¹⁴C]oleic, [¹⁴C]ricinoleic, or [¹⁴C]vernolic acid. Dioleoyl-PC that was labeled in the *sn*-1 position was synthesized from *sn*-1-[¹⁴C]oleoyl-lyso-PC and unlabeled oleic acid as described in (14). *Sn*-1-oleoyl-*sn*-2-[¹⁴C]ricinoleoyl-DAG was synthesized from PC by the action of phospholipase C type XI from *B. Cereus* (Sigma Chemical Co.) as described in (15). Monovernoloyl- and divernoleoyl-DAG were synthesized from TAG extracted from seeds of 15 *Euphorbia lagascae*, using the TAG-lipase (*Rizhopus arrhizus*, Sigma Chemical Co.) as previously described (16). Monoricinoleoyl-TAG was synthesized according to the same method using TAG extracted from Castor bean.

Lipid analysis. Total lipid composition of yeast were determined from cells 20 harvested from a 40 ml liquid culture, broken in a glass-bead shaker and extracted into chloroform as described by Bligh and Dyer (17), and then separated by thin layer chromatography in hexane/diethylether/acetic acid (80:20:1) using pre-coated silica gel 60 plates (Merck). The lipid areas were located by brief exposure to I₂ vapors and identified by means of appropriate 25 standards. Polar lipids, sterol-esters and triacylglycerols, as well as the remaining minor lipid classes, referred to as other lipids, were excised from the plates. Fatty acid methylesters were prepared by heating the dry excised material at 85 °C for 60 min in 2% (v/v) sulfuric acid in dry methanol. The methyl esters were extracted with hexane and analyzed by GLC through a 50 m 30 x 0.32 mm CP-Wax58-CB fused-silica column (Chrompack), with methylheptadecanoic acid as an internal standard. The fatty acid content of

each fraction was quantified and used to calculate the relative amount of each lipid class. In order to determine the total lipid content, 3 ml aliquots from yeast cultures were harvested by centrifugation and the resulting pellets were washed with distilled water and lyophilized. The weight of the dried cells was
5 determined and the fatty acid content was quantified by GLC-analyses after conversion to methylesters as described above. The lipid content was then calculated as nmol fatty acid (FA) per mg dry weight yeast.

Enzyme assays. Aliquots of crude microsomal fractions (corresponding to
10 10 nmol of microsomal PC) from developing plant seeds or yeast cells were lyophilized over night. ¹⁴C-Labeled substrate lipids dissolved in benzene were then added to the dried microsomes. The benzene was evaporated under a stream of N₂, leaving the lipids in direct contact with the membranes, and 0.1 ml of 50 mM potassium phosphate (pH 7.2) was added. The suspension was
15 thoroughly mixed and incubated at 30°C for the time period indicated, up to 90 min. Lipids were extracted from the reaction mixture using chloroform and separated by thin layer chromatography in hexane/diethylether/acetic acid (35:70:1.5) using silica gel 60 plates (Merck). The radioactive lipids were visualized and quantified on the plates by electronic autoradiography (Instant
20 Imager, Packard, US).

Yeast cultivation. Yeast cells were grown at 28°C on a rotatory shaker in liquid YPD medium (1% yeast extract, 2% peptone, 2% glucose), synthetic medium (18) containing 2% (v/v) glycerol and 2% (v/v) ethanol, or minimal
25 medium (19) containing 16 g/l of glycerol.

The instant invention is further characterized by the following examples which are not limiting:

30 Acyl-CoA-independent synthesis of TAG by oil seed microsomes. A large number of unusual fatty acids can be found in oil seeds (20). Many of these

fatty acids, such as ricinoleic (21) and vernolic acids (22), are synthesized using phosphatidylcholin (PC) with oleoyl or linoleoyl groups esterified to the *sn*-2 position, respectively, as the immediate precursor. However, even though PC can be a substrate for unusual fatty acid synthesis and is the major membrane lipids in seeds, unusual fatty acids are rarely found in the membranes. Instead, they are mainly incorporated into the TAG. A mechanism for efficient and selective transfer of these unusual acyl groups from PC into TAG must therefore exist in oil seeds that accumulate such unusual fatty acids.

This transfer reaction was biochemically characterized in seeds from castor bean (*Ricinus communis*) and *Crepis palaestina*, plants which accumulate high levels of ricinoleic and vernolic acid, respectively, and sunflower (*Helianthus annuus*), a plant which has only common fatty acids in its seed oil. Crude microsomal fractions from developing seeds were incubated with PC having ¹⁴C-labeled oleoyl, ricinoleoyl or vernoloyl groups at the *sn*-2 position. After the incubation, lipids were extracted and analyzed by thin layer chromatography. We found that the amount of radioactivity that was incorporated into the neutral lipid fraction increased linearly over a period of 4 hours (data not shown). The distribution of [¹⁴C]acyl groups within the neutral lipid fraction was analyzed after 80 min (Fig. 1). Interestingly the amount and distribution of radioactivity between different neutral lipids were strongly dependent both on the plant species and on the type of [¹⁴C]acyl chain. Thus, sunflower microsomes incorporated most of the label into DAG, regardless of the type of [¹⁴C]acyl group. In contrast, *R. communis* microsomes preferentially incorporated [¹⁴C]ricinoleoyl and [¹⁴C]vernoloyl groups into TAG, while [¹⁴C]oleyl groups mostly were found in DAG. *C. palaestina* microsomes, finally, incorporated only [¹⁴C]vernoloyl groups into TAG, with [¹⁴C]ricinoleyl groups being found mostly as free fatty acids, and [¹⁴C]oleyl groups in DAG. This shows that the high *in vivo* levels of ricinoleic acid and vernolic acid in the TAG pool of *R. communis* and *C. palaestina*, respectively, can be explained by an efficient and selective transfer of the corresponding acyl groups from PC to TAG in these organisms.

The in-vitro synthesis of triacylglycerols in microsomal preparations of developing castor bean is summarized in table 1.

PDAT: a novel enzyme that catalyzes acyl-CoA independent synthesis of TAG. It was investigated if DAG could serve both as an acyl donor as well as an acyl acceptor in the reactions catalyzed by the oil seed microsomes. Thererfore, unlabeled divernoloyl-DAG was incubated with either *sn*-1-oleoyl-*sn*-2-[¹⁴C]ricinoleoyl-DAG or *sn*-1-oleoyl-*sn*-2-[¹⁴C]ricinoleoyl-PC in the presence of *R. communis* microsomes. The synthesis of TAG molecules containing both [¹⁴C]ricinoleoyl and vernoloyl groups was 5 fold higher when [¹⁴C]ricinoleoyl-PC served as acyl donor as compared to [¹⁴C]ricinoleoyl-DAG (fig.1B). These data strongly suggests that PC is the immediate acyl donor and DAG the acyl acceptor in the acyl-CoA-independent formation of TAG by oil seed microsomes. Therefore, this reaction is catalyzed by a new enzyme which we call phospholipid : diacylglycerol acyltransferase (PDAT).

PDAT activity in yeast microsomes. Wild type yeast cells were cultivated under conditions where TAG synthesis is induced. Microsomal membranes were prepared from these cells and incubated with *sn*-2-[¹⁴C]-ricinoleoyl-PC and DAG and the ¹⁴C-labeled products formed were analyzed. The PC-derived [¹⁴C]ricinoleoyl groups within the neutral lipid fraction mainly were found in free fatty acids or TAG, and also that the amount of TAG synthesized was dependent on the amount of DAG that was added to the reaction (Fig.2). The *in vitro* synthesis of TAG containing both ricinoleoyl and vernoloyl groups, a TAG species not present *in vivo*, from exogenous added *sn*-2-[¹⁴C]ricinoleoyl-PC and unlabelled vernoloyl-DAG (Fig. 2, lane 3) clearly demonstrates the existence of an acyl-CoA-independent synthesis of TAG involving PC and DAG as substrates in yeast microsomal membranes. Consequently, TAG synthesis in yeast can be catalyzed by an enzyme similar to the PDAT found in plants.

The PDAT encoding gene in yeast.

A gene in the yeast genome (YNR008w) is known, but nothing is known about the function of YNR008w, except that the gene is not essential for growth under normal circumstances. Microsomal membranes were prepared

5 from the yeast strain FVKT004-04C(AL) (8) in which this gene with unknown function had been disrupted. PDAT activity in the microsomes were assayed using PC with radiolabelled fatty acids at the sn-2 position. The activity was found to be completely absent in the disruption strain (Fig. 2 lane 4).

10 Significantly, the activity could be partially restored by the presence of YNR008w on the single copy plasmid pUS1 (Fig. 2 lane 5). Moreover, acyl groups of phosphatidylethanolamine (PE) were efficiently incorporated into

TAG by microsomes from the wild type strain whereas no incorporation occurred from this substrate in the mutant strain. This shows that YNR008w encodes a yeast PDAT which catalyzes the transfer of an acyl group from the sn-2 position of phospholipids to DAG, thus forming TAG. It should be noted that no cholesterol esters were formed from radioactive PC even in incubations with added ergosterols, nor were the amount of radioactive free fatty acids formed from PC affected by disruption of the YNR008w gene. This demonstrates that yeast PDAT do not have cholesterol ester synthesising or

15 20 phospholipase activities.

RECEIVED IN BUREAU 10 JUNE 1997
ACCEPTED 10 JULY 1997
PUBLISHED 10 AUGUST 1997

Increased TAG content in yeast cells that overexpress PDAT. The effect of overexpressing the PDAT-encoding gene was studied by transforming a wild type yeast strain with the pUS4 plasmid in which the gene is expressed from the galactose-induced *GAL1:TPK2* promoter. Cells containing the empty expression vector were used as a control. The cells were grown in synthetic glycerol-ethanol medium, and expression of the gene was induced after either 2 hours (early log phase) or 25 hours (stationary phase) by the addition of galactose. The cells were then incubated for another 21 hours, after which they were harvested and assays were performed. We found that overexpression of PDAT had no significant effect on the growth rate as determined by the optical density. However, the total lipid content, measured as total μ mol fatty acids per mg yeast dry weight, was 47% (log phase) or 29% (stationary phase) higher in the PDAT overexpressing strain than in the control. Furthermore, the polar lipid and sterolester content was unaffected by overexpression of PDAT. Instead, the elevated lipid content in these cells is entirely due to an increased TAG content (Fig. 3A,B). Thus, the amount of TAG was increased by 2-fold in PDAT overexpressing early log phase cells and by 40% in stationary phase cells. It is interesting to note that a significant increase in the TAG content was achieved by overexpressing PDAT even under conditions (*i.e.* in stationary phase) where DAGAT is induced and thus contributes significantly to TAG synthesis. *In vitro* PDAT activity assayed in microsomes from the PDAT overexpressing strain was 7-fold higher than in the control strain, a finding which is consistent with the increased levels of TAG that we observed *in vivo* (Fig. 3C). These results clearly demonstrate the potential use of the PDAT gene in increasing the oil content in transgenic organisms.

Substrate specificity of yeast PDAT. The substrate specificity of yeast PDAT was analyzed using microsomes prepared from the PDAT overexpressing strain (see Fig. 4). The rate of TAG synthesis, under conditions given in figure 4 with di-oleoyl-PC as the acyl-donor, was 0.15 nmol per min

and mg protein. With both oleoyl groups of PC labeled it was possible, under the given assay conditions, to detect the transfer of 11 pmol/min of [¹⁴C]oleoyl chain into TAG and the formation of 15 pmol/min of lyso-PC. In microsomes from the PDAT-deficient strain, no TAG at all and only trace amounts of lyso-

- 5 PC was detected, strongly suggesting that yeast PDAT catalyses the formation of equimolar amounts of TAG and lyso-PC when supplied with PC and DAG as substrates. The fact that somewhat more lyso-PC than TAG is formed can be explained by the presence of a phospholipase in yeast microsomes, which produces lyso-PC and unesterified fatty acids from PC (data not shown).

10

- The specificity of yeast PDAT for different acyl group positions was investigated by incubating the microsomes with di-oleoyl-PC carrying a [¹⁴C]acyl group either at the *sn*-1 position (Fig. 4A bar 2) or the *sn*-2 position (Fig. 4A bar 3). We found that the major ¹⁴C-labeled product formed in the former case was lyso-PC, and in the latter case TAG. We conclude that yeast PDAT has a specificity for the transfer of acyl groups from the *sn*-2 position of the phospholipid to DAG, thus forming *sn*-1-lyso-PC and TAG. Under the given assay conditions, trace amounts of ¹⁴C-labelled DAG is formed from the *sn*-1 labeled PC by the reversible action of a CDP-choline : choline phosphotransferase (data not shown). This labeled DAG can then be further converted into TAG by the PDAT activity. It is therefore not possible to distinguish whether the minor amounts of labeled TAG that is formed in the presence of di-oleoyl-PC carrying a [¹⁴C]acyl group in the *sn*-1 position, is synthesized directly from the *sn*-1-labeled PC by a PDAT that also can act on the *sn*-1 position, or if it is first converted to *sn*-1-labeled DAG and then acylated by a PDAT with strict selectivity for the transfer of acyl groups at the *sn*-2 position of PC. Taken together, this shows that the PDAT encoded by YNR008w catalyses an acyl transfer from the *sn*-2 position of PC to DAG, thus causing the formation of TAG and lyso-PC.

20

30

The substrate specificity of yeast PDAT was further analyzed with respect to the headgroup of the acyl donor, the acyl group transferred and the acyl chains of the acceptor DAG molecule. The two major membrane lipids of *S. cerevisiae* are PC and PE, and as shown in Fig. 4B (bars 1 and 2), dioleoyl-PE
 5 is nearly 4-fold more efficient than dioleoyl-PC as acyl donor in the PDAT-catalyzed reaction. Moreover, the rate of acyl transfer is strongly dependent on the type of acyl group that is transferred. Thus, a ricinoleoyl group at the *sn*-2 position of PC is 2.5 times more efficiently transferred into TAG than an oleoyl group in the same position (Fig. 4B bars 1 and 3). In contrast, yeast PDAT has
 10 no preference for the transfer of vernoloyl groups over oleoyl groups (Fig. 4B bars 1 and 4). The acyl chain of the acceptor DAG molecule also affects the efficiency of the reaction. Thus, DAG with a ricinoleoyl or a vernoloyl group is a more efficient acyl acceptor than dioleoyl-DAG (Fig. 4B bars 1, 5 and 6). Taken together, these results clearly show that the efficiency of the PDAT-catalyzed
 15 acyl transfer is strongly dependent on the properties of the substrate lipids.

PDAT genes. Nucleotide and amino acid sequences of several PDAT genes are given as SEQ ID No. 1 through 15. Further provisional and/or partial sequences are given as SEQ ID NO 1a through 5a and 1b through 11b,
 20 respectively. One of the *Arabidopsis* genomic sequences (SEQ ID NO. 4) identified an *Arabidopsis* EST cDNA clone; T04806. This cDNA clone was fully characterised and the nucleotide sequence is given as SEQ ID NO. 5. Based on the sequence homology of the T04806 cDNA and the *Arabidopsis thaliana* genomic DNA sequence (SEQ ID NO 4) it is apparent that an additional A is
 25 present at position 417 in the cDNA clone (data not shown). Excluding this nucleotide would give the amino acid sequence depicted in SEQ ID NO. 12.

Increased TAG content in seeds of Arabidopsis thaliana that express the yeast PDAT. For the expression of the yeast pdat gene in *Arabidopsis thaliana* an EcoRI fragment from the pBluescript-pdat was cloned together with napin promotor (26) into the vector pGPTV-KAN (27). A plasmid (pGNapPDAT)

having the yeast PDAT gene in the correct orientation was identified and transformed into *Agrobacterium tumefaciens*. These bacteria were used to transform *Arabidopsis thaliana* columbia (C-24) plants using the root transformation method (28). Plants transformed with an empty vector were
5 used as controls.

First generation seeds (T1) were harvested and germinated on kanamycin containing medium. Second generation seeds (T2) were pooled from individual plants and their fatty acid contents analysed by quantification of their methyl esters by gas liquid chromatography after methylation of the seeds with 2% sulphuric acid in methanol at 85 °C for 1,5 hours. Quantification was done with heptadecanoic acid methyl esters as internal standard.
10

From the transformation with pGNapPDAT one T1 plant (26-14) gave raise to seven T2 plants of which 3 plants yielded seeds with statistically (in a mean difference two-sided test) higher oil content than seeds from T2 plants
15 generated from T1 plant 32-4 transformed with an empty vector (table 2).

PUBLISHED IN THE USA

Claims

1. An enzyme catalysing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.
5
2. An enzyme according to claim 1, comprising an amino acid sequence as set forth in SEQ ID No. 2 or a functional fragment, derivate, allele, homologue or isoenzyme thereof.
10
3. An enzyme according to claims 1 or 2 designated as phospholipid:diacylglycerol acyltransferase (PDAT).
15
4. An enzyme according to claims 1 to 3, comprising an amino acid sequence as set forth in SEQ ID No. 1a, 2b or 5a or a functional fragment, derivate, allele, homologue or isoenzyme thereof.
5. An enzyme according to claims 1 to 4, comprising an amino acid sequence selected from the group consisting of sequences as set forth in SEQ ID No. 2a, 3a, 5b, 6, 7b, 8, 13, 14, 15 or a functional fragment, derivate, allele, homologue or isoenzyme thereof.
20
6. An enzyme according to claims 1 to 5, comprising an amino acid sequence encoded through a nucleotide sequence, a portion, derivate, allele or homologue thereof selected from the group consisting of sequences as set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b, 12 or a functional fragment, derivate, allele, homologue or isoenzyme of the enzyme encoding amino acid sequence.
25
7. A nucleotide sequence encoding an enzyme catalysing in an acyl-CoA-independent reaction the transfer of fatty acids from phospholipids to
30

diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.

8. A nucleotide sequence according to claim 7 encoding an enzyme
5 designated as phospholipid:diacylglycerol acyltransferase (PDAT).
9. A nucleotide sequence according to claims 7 or 8, selected from the group consisting of sequences as set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b,
10 9b, 10, 10b or 11 or a portion, derivate, allele or homologue thereof.
10. A partial nucleotide sequence corresponding to a fulllength nucleotide sequence according to claims 7 to 9, selected from the group consisting of sequences as set forth in SEQ ID No. 5, 5b, 6b, 7, 8b, 9, 11b or 12 or a portion, derivate, allele or homologue thereof.
15
11. A nucleotide sequence according to claims 7 to 10, comprising a nucleotide sequence which is at least 40% homologous to a nucleotide sequence selected form the group consisting of those sequences set forth in SEQ ID No. 1, 1b, 3, 3b, 4, 4a, 4b, 5, 5b, 6b, 7, 8b, 9, 9b, 10, 10b, 11, 11b or 12.
20
12. A gene construct comprising a nucleotide sequence according to claims 7 to 11 operably linked to a heterologous nucleic acid.
25
13. A vector comprising a nucleotide sequence according to claims 7 to 11 or a gene construct according to claim 12.
14. A vector according to claim 13, which is an expression vector.
15. A vector according to claims 13 or 14, further comprising a selectable marker gene and/or nucleotide sequences for the replication in a host cell or the integration into the genome of the host cell.
30

16. A transgenic cell or organism containing a nucleotide sequence according to claims 7 to 11 and/or a gene construct according to claim 12 and/or a vector according to claims 13 to 15.
- 5
17. A transgenic cell or organism according to claim 16 which is an eucaryotic cell or organism.
- 10
18. A transgenic cell or organism according to claims 16 or 17 which is a yeast cell or a plant cell or a plant.
19. A transgenic cell or organism according to claims 16 to 18 having an altered biosynthetic pathway for the production of triacylglycerol.
- 15
20. A transgenic cell or organism according to claims 16 to 19 having an altered oil content.
21. A transgenic cell or organism according to claims 16 to 20 wherein the activity of PDAT is altered.
- 20
22. A transgenic cell or organism according to claims 16 to 21 wherein the altered activity of PDAT is characterized by an alteration in gene expression, catalytic activity and/or regulation of activity of the enzyme.
- 25
23. A transgenic cell or organism according to claims 16 to 22 wherein the altered biosynthetic pathway for the production of triacylglycerol is characterized by the prevention of accumulation of undesirable fatty acids in the membrane lipids.
- 30
24. A process for the production of triacylglycerol, comprising growing a transgenic cell or organism according to claims 16 to 23 under conditions

whereby the said nucleotide sequence according to claims 7 to 11 is expressed and whereby the said transgenic cells comprising an enzyme according to claims 1 to 6 catalysing the transfer of fatty acids from phospholipids to diacylglycerol forming triacylglycerol.

5

25. Triacylglycerols produced by a process according to claim 24.

26. Use of a nucleotide sequence according to claims 7 to 11 and/or an enzyme according to claims 1 to 6 for the production of triacylglycerol and/or triacylglycerols with uncommon fatty acids.

10

27. Use of a nucleotide sequence according to claims 7 to 11 and/or an enzyme according to claims 1 to 6 for the transformation of any cell or organism in order to be expressed in this cell or organism and result in an altered, preferably increased oil content of this cell or organism.

15

DRAFT DRAFT DRAFT DRAFT DRAFT

Abstract of the Disclosure

The present invention relates to the isolation, identification and characterization
5 of nucleotide sequences encoding an enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol, to the said enzymes and a process for the production of triacylglycerols.

10

DRAFT DRAFT DRAFT DRAFT DRAFT

References cited in the description:

1. Bell, R. M. & Coleman, R. A. (1980) *Annu. Rev. Biochem.* **49**, 459-487.
2. Stymne, S. & Stobart, K. (1987) in *The biochemistry of plants: a comprehensive treatise*, Vol. 9, ed. Stumpf, P. K. (Academic Press, New York), pp. 175-214.
- 5 3. Cases, S. et al. (1998) *Proc. Natl. Acad. Sci. U S A* **95**, 13018-13023.
4. Hobbs, D. H., Lu, C. & Hills, M. J. (1999) *FEBS Lett.* **452**, 145-9
- 5 5. Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G. & Taylor, D. C. (1999)
10 *Plant J.* **19**, 645-653.
6. Lardizabal, K., Hawkins, D., Mai, J., & Wagner, N. (1999) Abstract presented at the Biochem. Mol. Plant Fatty Acids Glycerolipids Symposium, South Lake Tahoe, USA.
7. Thomas, B. J. & Rothstein, R. (1989) *Cell* **56**, 619-630.
- 15 8. Entian, K.-D. & Kötter, P. (1998) *Meth. Microbiol.* **26**, 431-449.
9. Kern, L., de Montigny, J., Jund, R. & Lacroute, F. (1990) *Gene* **88**, 149-157.
10. Ronne, H., Carlberg, M., Hu, G.-Z. & Nehlin, J. O. (1991) *Mol. Cell. Biol.* **11**, 4876-4884.
11. Stobart, K. & Stymne, S. (1990) in *Method in Plant Biochemistry*, vol 4,
20 eds. Harwood, J. L. & Bowyer, J. R. (Academic press, London), pp. 19-46.
12. Bafor, M., Smith, M. A., Jonsson, L., Stobrt, A. K. & Stymne, S. (1991)
Biochem. J. **280**, 507-514.
13. Banas, A., Johansson, I. & Stymne, S. (1992) *Plant Science* **84**, 137-144.
14. Kanda, P. & Wells, M. A. (1981) *J. Lipid. Res.* **22**, 877-879.
- 25 15. Ståhl, U., Ek, B. & Stymne, S. (1998) *Plant Physiol.* **117**, 197-205.
16. Stobart, K., Mancha, M. & Lenman M, Dahlqvist, A. & Stymne, S. (1997)
Planta **203**, 58-66.
17. Bligh, E. G. & Dyer, W. J. (1959) *Can. J. Biochem. Physiol.* **37**, 911-917.
18. Sherman, F., Fink, G. R. & Hicks, J. B. (1986) in *Laboratory Course Manual
30 for Methods in Yeast Genetics* (Cold Spring Harbor Laboratory)
19. Meesters, P. A. E. P., Huijberts, G. N. M. and Eggink, G. (1996) *Appl.
Microbiol. Biotechnol.* **45**, 575-579.
20. van de Loo, F. J., Fox, B. G. & Sommerville, C. (1993), in *Lipid metabolism
in plants*, ed. Moore, T. S. (CRC Press, Inc.), pp. 91-126.
- 35 21. van de Loo, F. J., Broun, P., Turner, S. & Sommerville, S. (1995) *Proc. Natl.*

- Acad. Sci. U S A* **95**, 6743-6747.
22. Lee, M., Lenman, M., Banas, A., Bafor, M., Singh, S., Schweizer, M., Nilsson, R., Liljenberg, C., Dahlqvist, A., Gummesson, P-O., Sjödahl, S., Green, A., and Stymne, S. (1998) *Science* **280**, 915-918.
- 5 23. Glomset, J. A. (1968) *J. Lipid Res.* **9**, 155-167.
24. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucl. Acids Res.* **24**, 4876-4882.
25. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406-425.
- 10 26. Stålberg, K., Ellerström, M., Josefsson, L., & Rask, L. (1993) *Plant Mol. Biol.* **23**, 671
27. Becker, D., Kemper, E., Schell, J., Masterson, R. (1992) *Plant Mol. Biol.* **20**, 1195
28. D. Valvekens, M. Van Montagu, and Van Lusbettens (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 5536

15

Description of Figures

FIG. 1.

Metabolism of ^{14}C -labeled PC into the neutral lipid fraction by plant microsomes. (A) Microsomes from developing seeds of sunflower, *R. communis* and *C. palaestina* were incubated for 80 min at 30°C with PC (8 nmol) having oleic acid in its *sn*-1 position, and either ^{14}C -labeled oleic, ricinoleic or vernolic acid in its *sn*-2 position. Radioactivity incorporated in TAG (open bars), DAG (solid bars), and unsterified fatty acids (hatched bars) was quantified using thin layer chromatography followed by electronic autoradiography, and is shown as percentage of added labeled substrate. (B) Synthesis *in vitro* of TAG carrying two vernoloyl and one $[^{14}\text{C}]$ ricinoleoyl group by microsomes from *R. communis*. The substrates added were unlabeled divernoloyl-DAG (5 nmol), together with either *sn*-1-oleoyl-*sn*-2- $[^{14}\text{C}]$ ricinoleoyl-DAG (0.4 nmol, 7700 dpm/nmol) or *sn*-1-oleoyl-*sn*-2- $[^{14}\text{C}]$ ricinoleoyl-PC (0.4 nmol, 7700 dpm/nmol). The microsomes were incubated with the substrates for 30 min at 30°C, after which samples were removed for lipid analysis as described in the section „general methods“. The data shown are the average of two experiments.

20

FIG. 2.

PDAT activity in yeast microsomes, as visualized by autoradiogram of neutral lipid products separated on TLC. Microsomal membranes (10 nmol of PC) from the wild type yeast strain FY1679 (lanes 1-3), a congeneric yeast strain (FVKT004-04C(AL)) that is disrupted for YNR008w (lane 4) or the same disruption strain transformed with the plasmid pUS1, containing the YNR008w gene behind its native promotor (lane 5), were assayed for PDAT activity. As substrates, we used 2 nmol *sn*-1-oleoyl-*sn*-2- $[^{14}\text{C}]$ ricinoleoyl-PC together with either 5 nmol of dioleoyl-DAG (lanes 2, 4 and 5) or *rac*-oleoyl-vernoleoyl-DAG (lane 3). The enzymatic assay and lipid analysis was performed as described in Materials and Methods. The cells were precultured for 20 h in liquid YPD

medium, harvested and re-suspended in an equal volume of minimal medium (19) containing 16 g/l glycerol. The cells were then grown for an additional 24 h prior to being harvested. Selection for the plasmid was maintained by growing the transformed cells in synthetic medium lacking uracil (18). Abbreviations: 1-

5 OH-TAG, monoricinoleoyl-TAG; 1-OH-1-ep-TAG, monoricinoleoyl-monovernoloyl-TAG; OH-FA, unesterified ricinoleic acid.

Fig. 3.

Lipid content (A,B) and PDAT activity (C) in PDAT overexpressing yeast cells.

10 The PDAT gene in the plasmid pUS4 was overexpressed from the galactose-induced *GAL1-TPK2* promotor in the wild type strain W303-1A (7). Its expression was induced after (A) 2 hours or (B) 25 hours of growth by the addition of 2% final concentration (w/v) of galactose. The cells were then incubated for another 22 hours before being harvested. The amount of lipids of

15 the harvested cells was determined by GLC-analysis of its fatty acid contents and is presented as μ mol fatty acids per mg dry weight in either TAG (open bar), polar lipids (hatched bar), sterol esters (solid bar) and other lipids (striped bar). The data shown are the mean values of results with three independent yeast cultures. (C) *In vitro* synthesis of TAG by microsomes prepared from

20 yeast cells containing either the empty vector (vector) or the PDAT plasmid (+ PDAT). The cells were grown as in Fig. 3A. The substrate lipids dioleoyl-DAG (2.5 nmol) and *sn*-1-oleoyl-*sn*-2-[¹⁴C]-oleoyl-PC (2 nmol) were added to aliquots of microsomes (10 nmol PC), which were then incubated for 10 min at 28 °C. The amount of label incorporated into TAG was quantified by electronic

25 autoradiography. The results shown are the mean values of two experiments.

FIG. 4.

Substrate specificity of yeast PDAT. The PDAT activity was assayed by incubating aliquots of lyophilized microsomes (10 nmol PC) with substrate lipids at 30°C for 10 min (panel A) or 90 min (panel B). Unlabeled DAG (2.5 nmol) was used as substrates together with different labeled phospholipids, as shown

in the figure. (A) *Sn*-position specificity of yeast PDAT regarding the acyl donor substrate. Dioleoyl-DAG together with either *sn*-1-[¹⁴C]oleoyl-*sn*-2-[¹⁴C]oleoyl-PC (di-[¹⁴C]-PC), *sn*-1-[¹⁴C]oleoyl-*sn*-2-oleoyl-PC (*sn*1-[¹⁴C]-PC) or *sn*-1-oleoyl-*sn*-2-[¹⁴C]oleoyl-PC (*sn*2-[¹⁴C]-PC). (B) Specificity of yeast PDAT regarding 5 phospholipid headgroup and of the acyl composition of the phospholipid as well as of the diacylglycerol. Dioleoyl-DAG together with either *sn*-1-oleoyl-*sn*-2-[¹⁴C]oleoyl-PC (oleoyl-PC), *sn*-1-oleoyl-*sn*-2-[¹⁴C]oleoyl-PE (oleoyl-PE), *sn*-1-oleoyl-*sn*-2-[¹⁴C]ricinoleoyl-PC (ricinoleoyl-PC) or *sn*-1-oleoyl-*sn*-2-[¹⁴C]vernoloyl-PC (vernoloyl-PC). In the experiments presented in the 2 bars to 10 the far right, monoricinoleoyl-DAG (ricinoleoyl-DAG or mono-vernoloyl-DAG (vernoloyl-DAG) were used together with *sn*-1-oleoyl-*sn*-2-[¹⁴C]-oleoyl-PC. The label that was incorporated into TAG (solid bars) and lyso-PC (LPC, open bars) 15 was quantified by electronic autoradiography. The results shown are the mean values of two experiments. The microsomes used were from W303-1A cells overexpressing the PDAT gene from the *GAL1-TPK2* promotor, as described in Fig. 3. The expression was induced at early stationary phase and the cells were harvested after an additional 24 h.

20 TAB.1:

In vitro synthesis of triacylglycerols in microsomal preparations of developing castor bean. Aliquots of microsomes (20 nmol PC) were lyophilised and substrate lipids were added in benzene solution: (A) 0.4 nmol [¹⁴C]-DAG (7760 dpm/nmol) and where indicated 1.6 nmol unlabelled DAG; (B) 0.4 nmol [¹⁴C]-DAG (7760 dpm/nmol) and 5 nmol unlabelled di-ricinoleoyl-PC and (C) 0.25 nmol [¹⁴C]-PC (4000 dpm/nmol) and 5 nmol unlabelled DAG. The benzene was 25 evaporated by N₂ and 0.1 ml of 50 mM potassium phosphate was added, thoroughly mixed and incubated at 30 °C for (A) 20 min.; (B) and (C) 30 min.. Assays were terminated by extraction of the lipids in chloroform. The lipids 30 were then separated by thin layer chromatography on silica gel 60 plates

(Merck; Darmstadt, Germany) in hexan/diethylether/acetic 35:70:1.5. The radioactive lipids were visualised and the radioactivity quantified on the plate by electronic autoradiography (Instant Imager, Packard, US). Results are presented as mean values of two experiments.

5

Radioactivity in different triacylglycerols (TAG) species formed. Abbreviations used: 1-OH-, mono-ricinoleoyl-; 2-OH, di-ricinoleoyl-; 3-OH-, triricinoleoyl; 1-OH-1-ver-, mono-ricinoleoyl-monovernoleoyl-; 1-OH-2-ver-, mono-ricinoleoyl-divernoleoyl-. Radiolabelled DAG and PC were prepared enzymatically. The radiolabelled ricinoleoyl group is attached at the sn-2-position of the lipid and unlabelled oleoyl group at the sn-1-position. Unlabelled DAG with vernoleoyl- or ricinoleoyl chains were prepared by the action of TAG lipase (6) on oil of *Euphorbia lagascae* or Castor bean, respectively. Synthetic di-ricinoleoyl-PC was kindly provided from Metapontum Agribios (Italy).

10

TAB.2:

Total fatty acids per mg of T2 seeds pooled from individual *Arabidopsis thaliana* plants transformed with yeast PDAT gene under the control of napin promotor (26-14) or transformed with empty vector (32-4).

* = statistical difference between control plants and PDAT transformed plants in a mean difference two-sided test at $\alpha = 5$.

Description of the SEQ ID:

- SEQ ID NO. 1: Genomic DNA sequence and suggested amino acid sequence of the *Saccharomyces cerevisiae* PDAT gene, YNR008w, with GenBank accession number Z71623 and Y13139, and with nucleotide ID number 1302481.
- SEQ ID NO. 2: The amino acid sequence of the suggested open reading frame YNR008w from *Saccharomyces cerevisiae*.
- 10 SEQ ID NO. 3: Genomic DNA sequence of the *Schizosaccharomyces pombe* gene SPBC77E.14.
- SEQ ID NO. 4: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with GenBank accession number AB006704.
- 15 SEQ ID NO. 5: Nucleotide sequence of the *Arabidopsis thaliana* cDNA clone with GenBank accession number T04806, and nucleotide ID number 315966.
- SEQ ID NO. 6: Predicted amino acid sequence of the *Arabidopsis thaliana* cDNA clone with GenBank accession number T04806.
- 20 SEQ ID NO. 7: Nucleotide and amino acid sequence of the *Zea mays* EST clone with GenBank accession number AI491339, and nucleotide ID number g4388167.
- 25 SEQ ID NO. 8: Predicted amino acid sequence of the *Zea mays* EST clone with GenBank accession number AI491339, and nucleotide ID number g4388167.
- SEQ ID NO. 9: DNA sequence of part of the *Neurospora crassa* EST clone W07G1, with GenBank accession number AI398644, and nucleotide ID number g4241729.

SEQ ID NO. 10: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with GenBank accession number AC004557.

5 *SEQ ID NO. 11:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with GenBank accession number AC003027.

SEQ ID NO. 12: DNA sequencce of part of the *Lycopersicon esculentum* cDNA clone with GenBank accession number AI486635.

10 *SEQ ID NO. 13:* Amino acid sequence of the *Schizosaccharomyces pombe* putative open reading frame CAA22887 of the *Schizosaccharomyces pombe* gene SPBC776.14.

15 *SEQ ID NO. 14:* Amino acid sequence of the *Arabidopsis thaliana* putative open reading frame AAC80628 derived from the *Arabidopsis thaliana* locus with GenBank accession number AC004557.

20 *SEQ ID NO 15:* Amino acid sequence of the *Arabidopsis thaliana* putative open reading frame AAD10668 derived from the *Arabidopsis thaliana* locus with GenBank accession number AC003027.

Further provisional and/or partial sequences are defined through the following SEQ IDs:

25 *SEQ ID NO. 1a:* The amino acid sequence of the yeast ORF YNR008w from *Saccharomyces cerevisiae*.

SEQ ID NO. 2a: Amino acid sequence of the region of the *Arabidopsis thaliana* genomic sequence (AC004557).

SEQ ID NO. 3a: Amino acid sequence of the region of the *Arabidopsis thaliana* genomic sequence (AB006704).

5 SEQ ID NO. 4a: The corresponding genomic DNA sequence and amino acid sequence of the yeast ORF YNROO8w from *Saccharomyces cerevisiae*.

SEQ ID NO. 5a: The amino acid sequence of the yeast ORF YNROO8w from *Saccharomyces cerevisiae* derived form the corresponding genomic DNA sequence.

10

SEQ ID NO. 1b: Genomic DNA sequence of the *Saccharomyces cerevisiae* PDAT gene, YNR008w, genebank nucleotide ID number 1302481, and the suggested YNR008w amino acid sequence.

15

SEQ ID NO. 2b: The suggested amino acid sequence of the yeast gene YNR008w from *Saccharomyces cerevisiae*.

20

SEQ ID NO. 3b: Genomic DNA sequence of the *Schizosaccharomyces pombe* gene SPBC776.14.

SEQ ID NO. 4b: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AB006704.

25

SEQ ID NO. 5b: Nucleotide sequence and the corresponding amino acid sequence of the *Arabidopsis thaliana* EST-clone with genebank accession number T04806, and ID number 315966.

30

SEQ ID NO. 6b: Nucleotide and amino acid sequence of the *Zea mays* cDNA clone with genebank ID number g4388167.

SEQ ID NO. 7b: Amino acid sequence of the *Zea mays* cDNA clone with genebank ID number g4388167.

5 *SEQ ID NO. 8b:* DNA sequence of part of the *Neurospora crassa* cDNA clone
WO7G1, ID number g4241729.

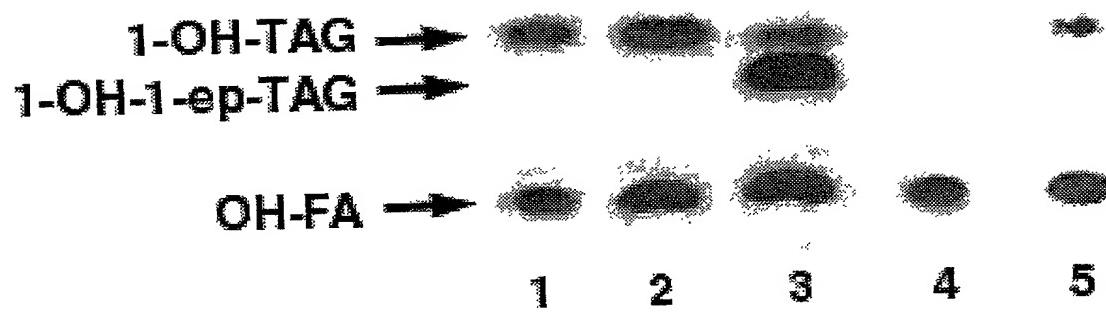
SEQ ID NO. 9b: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AC004557.

10 *SEQ ID NO. 10b:* Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AC003027.

SEQ ID NO. 11b: DNA sequence of part of the *Lycopersicon esculentum* cDNA clone with genebank accession number AI486635.

15

DOCUMENT ID: 00000000000000000000000000000000

Fig 2

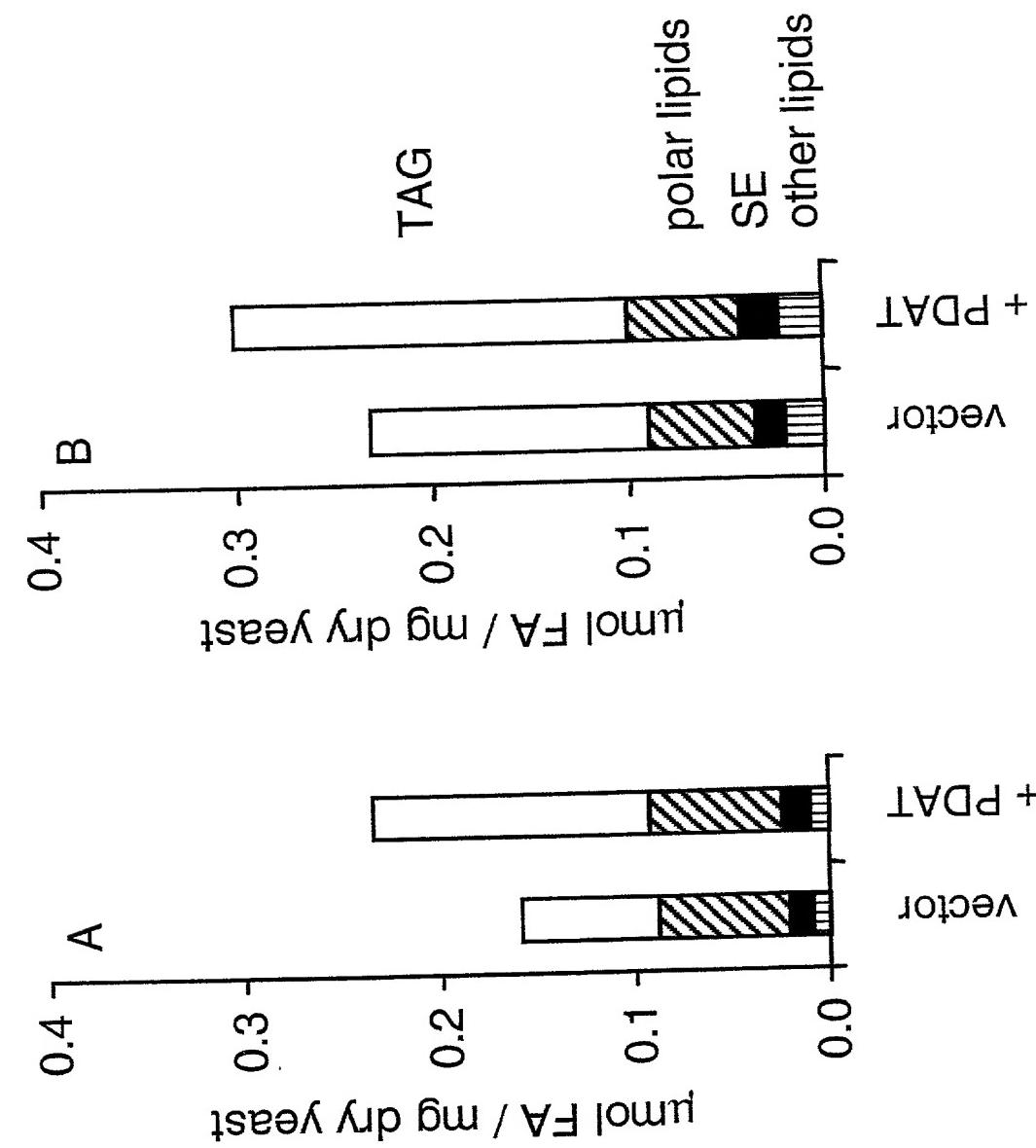
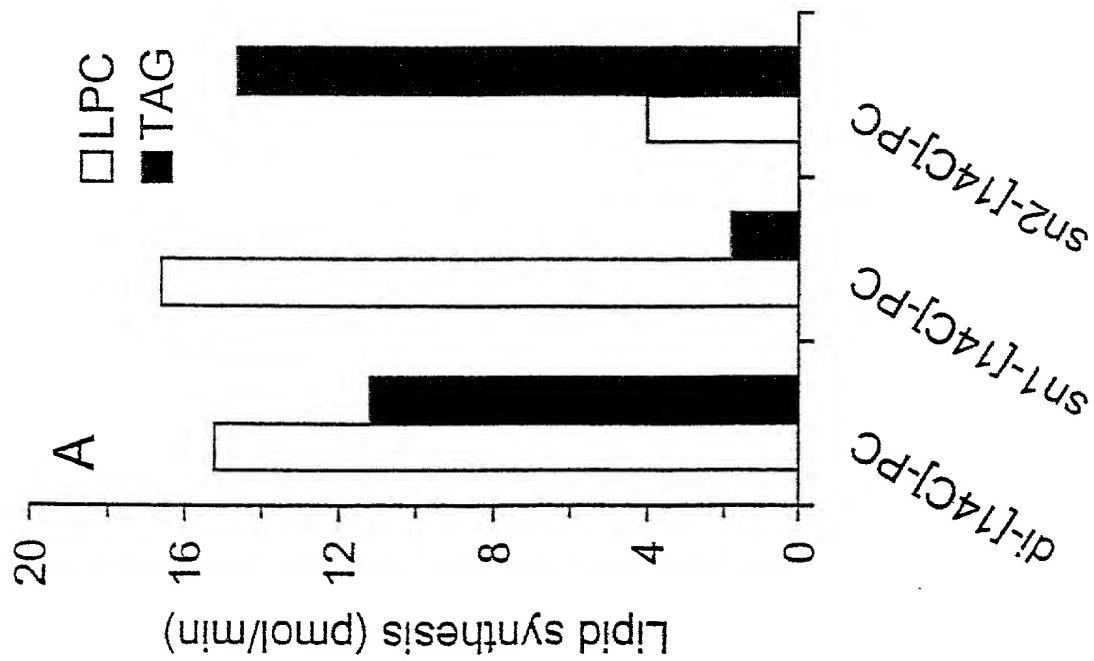


Fig. 3:

Fig.4:



Tables

Tab. 1:

Substrate added	$[^{14}\text{C}]\text{-lipid}^{(2)}$	unlabelled lipid ⁽²⁾	mol % of added $[^{14}\text{C}]\text{-acyl group in TAG}^{(1)}$				
			1-OH-TAG	2-OH-TAG	1-OH-1-ver-TAG	1-OH-2-ver-TAG	3-OH-TAG
A mono-[^{14}C]-ricinoleoyl-DAG	mono-ricinoleoyl-DAG	2,8	12,4	-	-	-	-
A mono-[^{14}C]-ricinoleoyl-DAG	mono-vernoleoyl-DAG	3,2	12,1	1,3	-	-	-
A mono-[^{14}C]-ricinoleoyl-DAG	di-vernoleoyl-DAG	4	10	0,5	1,2	-	-
A mono-[^{14}C]-ricinoleoyl-PC	di-ricinoleoyl-PC	0,3	24,8	-	-	-	-
B mono-[^{14}C]-ricinoleoyl-PC	none	6,8	8,0	-	-	4,7	-
C mono-[^{14}C]-ricinoleoyl-PC	di-oleoyl-DAG	8,6	9,8	-	-	-	5,0
C mono-[^{14}C]-ricinoleoyl-PC	mono-ricinoleoyl-DAG	5,7	16,7	-	-	-	1,9
C mono-[^{14}C]-ricinoleoyl-PC	di-ricinoleoyl-DAG	4,5	9,4	-	-	-	9,5
C mono-[^{14}C]-ricinoleoyl-PC	mono-vernoleoyl-DAG	6,0	11,5	10,9	0,5	0,5	7,4
C mono-[^{14}C]-ricinoleoyl-PC	di-vernoleoyl-DAG	6,7	10,8	1,1	8,4	8,4	6,8

Tab. 2:

T1 plant deviation	T2 plant number	nmol fatty acids per mg seed	standard
32-4	1	1277	±11 (n=2)
	4	1261	±63 (n=3)
	5	1369	±17 (n=3)
	6	1312	±53 (n=4)
	7	1197	±54 (n=5)
	8	1240	±78 (n=4)
	9	1283	±54 (n=5)
	10	1381	±35 (n=5)
26-14	1	1444	±110 (n=4)
	2	1617*	±109 (n=4)
	3	1374	±37 (n=2)
	5	1562*	±70 (n=4)
	6	1393	±77 (n=4)
	7	1433	±98 (n=4)
	8	1581*	±82 (n=4)

Sequence Listing

<210> 1

<211> 1986

<212> genomic DNA

<213> *Saccharomyces cerevisiae*

<221> CDS

<222> (1)..(1983)

<400> 1

atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct 48
 Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
 1 5 10 15

gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga 96
 Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
 20 25 30

aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt 144
 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
 35 40 45

att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg 192
 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
 50 55 60

aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg 240
 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
 65 70 75 80

att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt 288
 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
 85 90 95

ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt 336
 Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
 100 105 110

gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt 384
 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
 115 120 125

ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac 432
 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
 130 135 140

tac tcc aca tct tct gat gat ctc agt gaa aat ttt gcc gtt ggt Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly	145	150	155	160	480
aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val	165	170	175		528
atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile	180	185	190		576
gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp	195	200	205		624
gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp	210	215	220		672
ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn	225	230	235	240	
ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile	245	250	255		720
gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att Ala Gly Tyr Trp Ile Trp Asn Val Phe Gln Asn Leu Gly Val Ile	260	265	270		768
ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu	275	280	285		816
gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys	290	295	300		912
gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu	305	310	315	320	960
att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp	325	330	335		1008

gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt			1056
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val			
340	345	350	
aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc			1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly			
355	360	365	
gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc			1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr			
370	375	380	
att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca			1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser			
385	390	395	400
aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca			1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser			
405	410	415	
atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct			1296
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser			
420	425	430	
tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att			1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile			
435	440	445	
cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg			1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met			
450	455	460	
aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa			1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln			
465	470	475	480
aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa			1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu			
485	490	495	
gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg			1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met			
500	505	510	
gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac			1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr			
515	520	525	

ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat 1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
530 535 540

gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct 1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
545 550 555 560

gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca 1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
565 570 575

atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga 1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
580 585 590

att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat 1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
595 600 605

ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc 1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
610 615 620

gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat 1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
625 630 635 640

ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag 1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
645 650 655

atg ccc ttc cca atg taa 1986
Met Pro Phe Pro Met
660

<210> 2

<211> 661

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
1 5 10 15
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
20 25 30
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
35 40 45
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
50 55 60
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
65 70 75 80
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
85 90 95
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
100 105 110
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
115 120 125
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
130 135 140
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly
145 150 155 160
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
165 170 175
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile
180 185 190
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
195 200 205
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
210 215 220
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
225 230 235 240
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
245 250 255
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
260 265 270
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
275 280 285
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
290 295 300
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
305 310 315 320
Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
340 345 350
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
355 360 365
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
370 375 380
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
385 390 395 400
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
405 410 415
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
420 425 430
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
435 440 445
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
450 455 460
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
465 470 475 480
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
485 490 495
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
500 505 510
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
515 520 525
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
530 535 540
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
545 550 555 560
Val Phe Leu Thr Glu Asp Gly Thr Val Pro Leu Val Ala His Ser
565 570 575
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
580 585 590
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
595 600 605
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
610 615 620
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
625 630 635 640
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
645 650 655
Met Pro Phe Pro Met
660

<210> 3
<211> 2312
<212> genomic DNA
<213> Schizosaccharomyces pombe

<400> 3	
ATGGCGTCTT CCAAGAAGAG CAAAACTCAT AAGAAAAAGA AAGAAGTCAA	50
ATCTCCTATC GACTTACCAA ATTCAAAGAA ACCAACTCGC GCTTTGAGTG	100
AGCAACCTTC AGCGTCCGAA ACACAACCTG TTTCAAATAA ATCAAGAAAA	150
TCTAAATTG GAAAAGATT GAATTTATA TTGGCGCTA TTTTGGGAAT	200
ATGCGGTGCT TTTTTTTCTG CTGTTGGAGA CGACAATGCT GTTTTCGACC	250
CTGCTACGTT AGATAAATTG GGGAAATATGC TAGGCTCTTC AGACTTGTGTT	300
GATGACATTA AAGGATATTT ATCTTATAAT GTGTTAAGG ATGCACCTTT	350
TAATACGGAC AAGCCTTCGC AGTCTCCTAG CGGAAATGAA GTTCAAGTTG	400
GTCTTGATAT GTACAATGAG GGATATCGAA GTGACCATCC TGTTATTATG	450
GTTCCCTGGTG TTATCAGCTC AGGATTAGAA AGTGGTCTG TTAATAATTG	500
CTCGATTCTT TACCTTAGGA AACGTCTTG GGGTAGCTGG TCTATGCTGA	550
AGGCAATGTT CCTTGACAAG CAATGCTGGC TTGAACATTT AATGCTTGAT	600
AAAAAAACCG GCTTGGATCC GAAGGGAATT AACGTCGAG CAGCTCAGGG	650
GTTTGAAGCA GCTGATTTT TTATCACGGG CTATTGGATT TGGAGTAAAG	700
TAATTGAAAA CCTTGCTGCA ATTGGTTATG AGCCTAATAA CATGTTAAGT	750
GCTTCTTACG ATTGGCGGTT ATCATATGCA AATTAGAGG AACGTGATAA	800
ATATTTTCA AAGTAAAAAA TGTTCATGCA GTACAGCAAC ATTGTACATA	850
AGAAAAAGGT AGTGGTGTATT TCTCACTCCA TGGGTTACAA GGTTACGTAC	900
TATTTTTTA AGTGGGTTGAGCTGAGGGC TACGGAAATG GTGGACCGAC	950
TTGGGTTAAT GATCATATTG AAGCATTAT AAATGTGAGT CTCGATGGTT	1000
GTTTGACTAC GTTCTAACT TTTGAATAGA TATCGGGATC TTTGATTGGA	1050
GCACCCAAAA CAGTGGCAGC GCTTTATCG GGTGAAATGA AAGATACAGG	1100
TATTGTAATT ACATTAACAA TGTTAATATT TAATTTTTGC TAACCGTTTT	1150
AAGCTCAATT GAATCAGTT TCGGTCTATG GGTAAGCAAT AAATTGTTGA	1200
GATTTGTTAC TAATTTACTG TTTAGTTGG AAAAATTTT TTCCCGTTCT	1250
GAGGTATATT CAAAAATACA AATGTGCTCT ACTTTTCTA ACTTTAATA	1300
GAGAGCCATG ATGGTCGCA CTATGGGAGG AGTTAGTTCT ATGCTTCCTA	1350
AAGGAGGCCA TGTGTATGG GGAAATGCCA GTTGGGTAAG AAATATGTGC	1400
TGTTAATTG TTATTAATAT TTAGGCTCCA GATGATCTTA ATCAAACAAA	1450
TTTTTCCAAT GGTGCAATT TTCGATATAG AGAAGACATT GATAAGGACC	1500
ACGATGAATT TGACATAGAT GATGCATTAC AATTGTTAAA AAATGTTACA	1550
GATGACGATT TTAAAGTCAT GCTAGCGAAA AATTATTCCC ACGGTCTTGC	1600
TTGGACTGAA AAAGAAGTGT TAAAAAATAA CGAAATGCCG TCTAAATGGA	1650
TAATCCGCT AGAAGTAAGA ACATTAAGT TACTAAATTAA TACTAACCA	1700
AATAGACTAG TCTCCTTAT GCTCCTGATA TGAAAATTAA TTGCGTTCAC	1750
GGGGTCGAA ACCAAGTGA GAGAGTTAT TATTATACTA ATAATCCTGA	1800
GGGGCAACCT GTCATGATT CCTCGGTTAA TGATGGAACA AAAGTTGAAA	1850
ATGTGAGAGA ATTATGTTT CAAACATTCT ATTAACGTGTT TTATTAGGTT	1900
ATTGTTATGG ATGATGGTGA TGGAACCTTA CCAATATTAG CCCTTGGTTT	1950
GGTGTGCAAT AAAGTTGGC AAACAAAAAG GTTAATCCT GCTAATACAA	2000
GTATCACAA TTATGAAATC AAGCATGAAC CTGCTGCGTT TGATCTGAGA	2050
GGAGGACCTC GCTCGGCAGA ACACGTCGAT ATACCTGGAC ATTCAAGAGCT	2100
AAATGTATGT TCATTTTACCT TTACAAATTG CTATTACTAA CTCTTGAAAT	2150
AAGGAAATTA TTTAAAAGT TTCATCAGGC CATGGTGACT CGGTACCAAA	2200
CCGTTATATA TCAGATATCC AGTACGGACA TAAGTTTGTT AGATTGCAAT	2250
TAACTAACCA ACCAACAGG GAAATAATAA ATGAGATAAA TCTCGATAAA	2300
CCTAGAAATT AA	

<210> 4
<211> 3685
<212> genomic DNA
<213> Arabidopsis thaliana

<400> 4	
ATGCCCTTA TTCATCGGAA AAAGCCGACG GAGAAACCAT CGACGCCGCC	50
ATCTGAAGAG GTGGTGCACG ATGAGGATTG CCAAAAGAAA CCACACGAAT	100
CTTCAAATC CCACCATAAG AAATCGAACG GAGGAGGGAA GTGGTCGTGC	150
ATCGATTCTT GTTGTGGTT CATTGGGTGT GTGTGTGTA CCTGGTGGTT	200
TCTTCTCTTC CTTACAACG CAATGCCCTGC GAGCTCCCT CAGTATGTAA	250
CGGAGCGAAT CACGGGTCTT TTGCCTGACC CGCCCGGTGT TAAGCTCAA	300
AAAGAAGGTC TTAAGGCGAA ACATCCTGTT GTCTTCATTC CTGGGATTGT	350
CACCGGTGGG CTCGAGCTTT GGGAAAGGCAA ACAATGCGCT GATGGTTAT	400
TTAGAAAACG TTTGTGGGT GGAACTTTG GTGAAGTCTA CAAAAGGTGA	450
GCTCAACAAT TCTCACTCTT CCTTTATATT GGGATTGGA TTGGATCTGA	500
TGAGATCACG CACTTGTGTC TTCTTCAACA TCACCTCAAAC TTTAATTCCA	550
TGTTTGTCTG TCTACTCTT TACTTTTTT TTTTTTGAT GTGAAAGGCT	600
ATTTTCTAA GAGACTATTT CTGTATGTGT AAGGTAAGCG TTCCAAGGAC	650
GTAATTGGCT TGGACTATTT CTGTTGATT GTTAACCTTA GGATATAAAA	700
TAGCTGCCTT GGAATTCAA GTCATCTTAT TGCCAAATCT GTTGCTAGAC	750
ATGCCCTAGA GTCCGTTCAT AACAAAGTTAC TTCCCTTACT GTCGTTGCGT	800
GTAGATTAG CTTGTGTAG CGTATAATGA AGTAGTGT TATGTTTGT	850
TGGGAATAGA GAAGTTCTAA CTACATCTGT GGAAAGTGTG TTCAGGCTGT	900
GATAGAGGAC TGTGCTTTA TTATTCAACT ATGTATATGT GTAATTAAAG	950
CTAGTTCTT TTTGATCTTT CAGCTCAATG TGCTTTCTC AATTTTTTC	1000
TCAATTCAA AGTTTCACAT CGAGTTTATT CACATGTCTT GAATTTCGTC	1050
CATCCTCGTT CTGTTATCCA GCTTGAACCT CCTCCCGACC CTGCTATGGA	1100
TATATTAAAA AAAAAGTGTG TTGTGGGTG CATCTTTGTT ACGATCTGCA	1150
TCTTCTCTT TCGGCTCAGT GTTCATGTTT TTGCTATGGT AGAGATGGGC	1200
AATGTTATTG TTGATGGTAA CAGTGGTATA GTTGATAGTA TCTTAACCTAA	1250
TCAATTATCT CTTGATTCA GCCCTCTATG TTGGGTGGAA CACATGTAC	1300
TTGACAATGA AACTGGGTG GATCCAGCTG GTATTAGAGT TCGAGCTGTA	1350
TCAGGACTCG TGGCTGCTGA CTACTTGTCT CCTGGCTACT TTGTCGGGC	1400
AGTGCTGATT GCTAACCTTG CACATATTGG ATATGAAGAG AAAAATATGT	1450
ACATGGCTGC ATATGACTGG CGGCTTCGT TTOAGAACAC AGAGGTTCTT	1500
TTCTCATCGT TCTTCTATT ATTCTGTTCG ATGTTACGTT TCTTCTTCA	1550
TTACTTAAGG CTTAAATATG TTTCATGTTG AATTAATAGG TACGTGATCA	1600
GAECTCTAGC CGTATGAAAAA GTAATATAGA GTTGATGGGT TCTACCAACG	1650
GTGGAAAAAA AGCAGTTATA GTTCCCGCATT CCATGGGGGT CTGTTATTTT	1700
CTACATTATA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGCGGGGG	1750
TGGGCCAGAT TGGTGTGCAA AGTATATTAA GGCGGTGATG AACATTGGTG	1800
GACCATTCT TGGTGTTCGA AAAGCTGTTG CAGGGCTTTT CTCTGCTGAA	1850
GCAAAGGATG TTGCACTTGC CAGGTATGTA ATATCTGCTT ATACTTTGA	1900
TGATCAGAAC CTTGGCTCTG GAACTCAAAG TTATTCTACT AAATATCAAT	1950
TCTAATAACA TTGCTATATT ATCGCTGCAA CTGACATTGG TTGATTATTT	2000
TTGCTGCTTA TGTAACGTGAA ACTCTCTTGA GATTAGACAA ATGATGAATT	2050
GATAATTCTT ACCGATCTGC CTGTGATGAC CAGTTTCTTA GCTTCGACGA	2100
TAACATTGTG CATACTGTCT TTTGGAGGGC ATTGAATTTC GCTATGGAAA	2150
GCGCTGGACC TTCCATGCTT GCATTCTTA CCAATTAGCG TTATTCTGCT	2200
TCTTCAATT TTCTTGTATA TGCATCTATG GTCTTTTATT TCTTCTTAAT	2250
TAAAGACTCG TTGGATTAGT TGCTCTATTA GTCACTGGT TCCTTAATAT	2300
AGAACCTTAC TTTCTTCGAA AATTGCAAG CGATTGCCCC AGGATTCTTA	2350
GACACCGATA TATTTAGACT TCAGACCTTG CAGCATGTA TGAGAATGAC	2400
ACGCACATGG GACTCAACAA TGTCTATGTT ACCGAAGGGGA GGTGACACGA	2450
TATGGGGCGG GCTTGATTGG TCACCGGAGA AAGGCCACAC CTGTTGTGGG	2500
AAAAAGCAAA AGAACAAACGA AACCTGTGGT GAAGCAGGTG AAAACGGAGT	2550
TTCCAAGAAA AGTCCCTGTTA ACTATGGAAG GATGATATCT TTTGGGAAAG	2600
AAAGTAGCAGA GGCTGCGCCA TCTGAGATTA ATAATATTGA TTTTCGAGTA	2650
AGGACATATA AATCATATA AACCTTGAC ATTGTGTGAT TGTATGATGA	2700
ATATCTGTAC ATTCTATCTG GTGAAGGGTG CTGTCACAGG TCAGAGTATC	2750
CCAAATCACA CCTGTCGTGA CGTGTGGACA GAGTACCATG ACATGGGAAT	2800
TGCTGGGATC AAAGCTATCG CTGAGTATAA GGTCTACACT GCTGGTGAAG	2850

CTATAGATCT ACTACATTAT GTTGCTCCTA AGATGATGGC GCGTGGTGC	2900
GCTCATTTCT CTTATGGAAT TGCTGATGAT TTGGATGACA CCAAGTATCA	2950
AGATCCCAA TACTGGTCAA ATCCGTTAGA GACAAAGTAA GTGATTTCTT	3000
GATTCCAACG GTATCCTTCG TCCTGATGCA TTATCAGTCT TTTTGTTC	3050
GGTCTTGTG GATATGGTTT TCAGCTAAA GCTTACAAAG CTGTTCTGA	3100
GCCTTTCTCA AAAAGGCTTG CTCAGTAATA TTGAGGTGCT AAAGTTGATA	3150
CATGTGACTC TTGCTTATAA ATCCTCCGTT TGGTTGTTC TGCTTTCTCA	3200
GATTACCGAA TGCTCCTGAG ATGGAAATCT ACTCATTATA CGGAGTGGGG	3250
ATACCAACGG AACGAGCATA CGTATACAAG CTTAACCGT CTCCCGACAG	3300
TTGCATCCCC TTTCAGATAT TCACTTCTGC TCACGGAGGAG GACGAAGATA	3350
GCTGTCTGAA AGCAGGAGTT TACAATGTGG ATGGGGATGA AACAGTACCC	3400
GTCCTAAGTG CCGGGTACAT GTGTGCAAAA GCGTGGCGTG GCAAGACAAG	3450
ATTCAACCT TCCGGAATCA AGACTTATAT AAGAGAATAC AATCACTCTC	3500
CGCCGGCTAA CCTGTTGGAA GGGCGCGGGA CGCAGAGTGG TGCCCATGTT	3550
GATATCATGG GAAACTTGC TTTGATCGAA GATATCATGA GGGTTGCCGC	3600
CGGAGGTAAC GGGCTGTATA TAGGACATGA CCAGGTCCAC TCTGGCATAT	3650
TTGAATGGTC GGAGCGTATT GACCTGAAGC TGTGA	3685

<210> 5
<211> 2427
<212> cDNA
<213> Arabidopsis thaliana

<400> 5	
AGAAACAGCT CTTGTCTCT CTCGACTGAT CTAACAATCC CTAATCTGTG	50
TTCTAAATTCTC CTGGACGAGA TTTGACAAAG TCCGTATAGC TTAACCTGGT	100
TTAATTTCAA GTGACAGATA TGCCCCTAT TCATCGGAAA AAGCCGACGG	150
AGAAACCATC GACGCCGCCA TCTGAAGAGG TGGTGCACGA TGAGGATTCG	200
CAAAAGAACAC CACACGAATC TTCCAAATCC CACCATAAGA AATCGAACGG	250
AGGAGGGAAG TGGTCGTGCA TCGATTCTTG TTGTTGGTTC ATTGGGTGTG	300
TGTGTGTAAC CTGGTGGTTT CTTCTCTCC TTTACAACGC AATGCCTGCG	350
AGCTTCCCTC AGTATGTAAC GGAGCGAATC ACGGGTCCCTT TGCCTGACCC	400
GCCCCGGTGT AAGCTAAAAA AAAGAAGGTC TTAAGGCAGA ACATCCTGTT	450
GTCTTCATTC CTGGGATTGT CACCGGTTGG CTCGAGCTTT GGGAAAGGCAA	500
ACAATGCGCT GATGGTTTAT TTAGAAAACG TTTGTTGGGT GGAACCTTTG	550
GTGAAGTCTA CAAAAGGCCCT CTATGTTGGG TGGAAACACAT GTCACTTGAC	600
AATGAAACATG GGTGGATCC AGCTGGTATT AGAGTTCGAG CTGTATCAGG	650
ACTCGTGGCT GCTGACTACT TTGCTCCTGG CTACTTTGTC TGGGCAGTGC	700
TGATTGCTAA CCTTGCACAT ATTGGATATG AAGAGAAAAA TATGTACATG	750
GCTGCATATG ACTGGCGGCT TTCGTTTCAG AACACAGAGG TACGTGATCA	800
GACTCTTAGC CGTATGAAAAA GTAATATAGA GTTGATGGTT TCTACCAACG	850
GTGGAAAAAA AGCAAGTTATA GTTCCGCATT CCATGGGGGT CTTGTATTTT	900
CTACATTAA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGCGGGGG	950
TGGGCCAGAT TGGTGTGCAA AGTATATTAA GGCGGTGATG AACATTGGTG	1000
GACCATTCTC TGGTGTTCCTA AAAGCTGTTG CAGGGCTTTT CTCTGCTGAA	1050
GCAAAGGATG TTGCAGTTGC CAGAGCGATT GCCCCAGGAT TCTTAGACAC	1100
CGATATATTT AGACTTCAGA CCTTGCAGCA TGTAATGAGA ATGACACGCA	1150
CATGGGACTC ACAATGTCT ATGTTACCGA AGGGAGGTGA CACGATATGG	1200
GGCGGGCTTG ATTGGTCACC GGAGAAAGGC CACACCTGTT GTGGGAAAAAA	1250
GCAAAGAAC AACGAAACCTT GTGGTGAAGC AGGTGAAAAC GGAGTTTCCA	1300
AGAAAAGTCC TGTAACTAT GGAAGGATGA TATCTTTGG GAAAGAAGTA	1350
GCAGAGGCTG CGCCATCTGA GATTAATAAT ATTGATTTC GAGGTGCTGT	1400
CAAAGGTCAAG AGTATCCCAA ATCACACCTG TCGTGACGTG TGGACAGAGT	1450
ACCATGACAT GGGAAATTGCT GGGATCAAAG CTATCGCTGA GTATAAGGTC	1500
TACACTGCTG GTGAAGCTAT AGATCTACTA CATTATGTTG CTCCCTAACAT	1550
GATGGCGCGT GGTGCCGCTC ATTTCTCTTA TGGAAATTGCT GATGATTG	1600
ATGACACCAA GTATCAAGAT CCCAAATACT GGTCAAATCC GTTAGAGACA	1650
AAATTACCGA ATGCTCCTGA GATGGAAATC TACTCATTAT ACGGAGTGGG	1700
GATACCAACG GAACGAGCAT ACGTATACAA GCTTAACCAAG TCTCCCGACA	1750
GTTGCATCCC CTTTCAGATA TTCACTTCTG CTCACGAGGA GGACGAAGAT	1800
AGCTGTCTGA AACGAGGAGT TTACAATGTG GATGGGGATG AAACAGTACC	1850
CGTCCTAAGT GCCGGGTACA TGTGTGCAAAG AGCGTGGCGT GGCAAGACAA	1900
GATTCAACCC TTCCCGGAATC AAGACTTATA TAAGAGAATA CAATCACTCT	1950
CCGCCGGCTA ACCTGTTGGA AGGGCGCGGG ACGCAGAGTG GTGCCCATGT	2000
TGATATCATG GGAAACTTTG CTTTGATCGA AGATATCATG AGGGTTGCCG	2050
CCGGAGGTTAA CGGGTCTGAT ATAGGACATG ACCAGGTCCA CTCTGGCATA	2100
TTTGAATGGT CGGAGCGTAT TGACCTGAAAG CTGTGAATAT CATGATCTCT	2150
TTAAGCTGTC CTGTCAGCTT ATGTGAATCC AATACTTTGA AAGAGAGATC	2200
ATCATCAATT CATCATCATC GTCATCATCA TGATGCTCAA CTCACAAAGA	2250
AGCCTGAGAA TGATACTTTG GTgCGAAATT CTCAATACCT CTTTAATATT	2300
CTTATTGAAAT GTAAATTATA CAATCCTATC TAATGTTTGA ACGATAACAC	2350
AAAACTTGCT GCNGCCATGT TTGTTGTCT TGTCAAAAGC ATCAATTG	2400
GGGTAAAAAA AAAAAAAA AAAAAAA	2427

<210> 6
<211> 671
<212> PRT
<213> Arabidopsis thaliana

<400> 6

MPLIHRKKPT EKPSTPPSEE VVHDEDSQKK PHESSKSHHK KSNGGGKWSC	50
IDSCCCWFFIGC VCVTWWFLLF LYNAMPASFP QYVTERITGP LPDPPGVKLK	100
KEGLKAKHPV VFIPGIVTGG LELWEGKQCA DGLFRKRLWG GTFGEVYKRP	150
LCWVEHMSLD NETGLDPAGI RVRAVSGLVA ADYFAPGYFV WAVLIANLAH	200
IGYEEKNMYM AAYDWRLSFQ NTEVRDQTLS RMKSNIELMV STNGGKKAVI	250
VPHSMGVLYF LHFMKWVEAP APLGGGGPD WCAKYIKAVM NIGGPFLGVP	300
KAVAGLFSAE AKDVAVARAI APGFLDDIF RLQLQHVMR MTRTDWDSTMS	350
MLPKGGDTIW GGLDWSPEKG HTCCGKKQKN NETCGEAGEN GVSKKSPVNY	400
GRMISFGKEV AEAAPSEINN IDFRGAVKGQ SIPNHTCRDV WTEYHDMGIA	450
GIKATAEYKV YTAGEAIDLL HYVAPKMMAR GAAHFSYGIN DDLDDETKYQD	500
PKYWSNPLET KLPNAPEMEI YSLYGVGIPT ERAYVYKLNQ SPDSCIPFQI	550
FTSAHEEDED SCLKAGVYNV DGDETPVPLS AGYMCAKAWR GKTRFNPNSGI	600
KTYIREYNHS PPANLLEGRG TQSGAHVDIM GNFALEIDIM RVAAGGNGSD	650
IGHDQVHSGI FEWSERIDLK L	671

<210> 7
<211> 643
<212> cDNA
<213> Zea mays

<221> CDS
<222> (1)..(402)

<400> 6

CGG GAG AAA ATA GCT GCT TTG AAG GGG GGT GTT TAC TTA GCC GAT GGT 48
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
1 5 10 15

GAT GAA ACT GTT CCA GTT CTT AGT GCG GGC TAC ATG TGT GCG AAA GGA 96
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly
20 25 30

TGG CGT GGC AAA ACT CGT TTC AGC CCT GCC GGC AGC AAG ACT TAC GTG 144
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val
35 40 45

AGA GAA TAC AGC CAT TCG CCA CCC TCT ACT CTC CTG GAA GGC AGG GGC 192
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly
50 55 60

ACC CAG AGC GGT GCA CAT GTT GAT ATA ATG GGG AAC TTT GCT CTA ATT 240
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile
65 70 75 80

GAG GAC GTC ATC AGA ATA GCT GCT GGG GCA ACC GGT GAG GAA ATT GGT 288
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly
85 90 95

GCC GAT CAG GTT TAT TCA GAT ATA TTC AAG TGG TCA GAG AAA ATC AAA 336
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys
100 105 110

TTG AAA TTG TAA CCTATGGAA GTTAAAGAAC TGCCGACCCG TTTATTGCGTTCC 391
Leu Lys Leu
115

AAAGTGTCT GCCTGAGTGC AACTCTGGAT TTTGCTAAA TATTGTAATT TTTCACGC 449
TTCATTCGTC CCTTGTCAA ATTTACATT GACAGGACGC CAATCGATA CGATGTTG 507
TACCGCTATT TTCAGCATTG TATATTAAC TGTACAGGTG TAAGTTGCAT TTGCCAGC 565
TGAAATTGTG TAGTCGTTT CTTTACGATT TAATANCAAG TGGCGGAGCA GTGCCCCA 623
AGCAAAAAAA AAAAAAAA 643

<210> 8
<211> 115
<212> PRT
<213> Zea mays

<400> 8
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
1 5 10 15

Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly
20 25 30

Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val
35 40 45

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly
50 55 60

Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile
65 70 75 80

Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly
85 90 95

Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys
100 105 110

Leu Lys Leu
115

<210> 9
<211> 616
<212> cDNA
<213> Neurospora crassa

<400> 9
ggtgtggcgaag acganggcgg aagttggagg ctaacgagaa tgacnctcg 50
agatggatct accctctaga gacacgacta ccntgcacc cagcctcaag 100
gtntacngtt tntatgggt a ggaagccgac ggagcgagcc tacatctatc 150
tggcgccccga tcccccggacg acaacgcacat ttttagatgac gatcgatacg 200
actttgactn agggcacaat tgaccacggt gtgattttgg g c g a a g g c g a 250
tggcacagtg aaccttatga gtttgggt a c t g t g c a a t a a g g g g t g g a 300
aaatgaagag atacaatcct g c g g g c t c a a a a t a a c c g t g g t c g a g a t g 350
ccgcatgaac cagaacgggtt caatccgaga ggagggccga atacggcgga 400
tcacgtggat attcttaggaa ggcagaatct aaacgagtac attcttaaag 450
tggcggcagg tcgaggcgat acaattgagg attttattac tagtaatatt 500
cttaaatatg tagaaaaggt tgaaattttat gaagagtaat taaaataccgc 550
acataggtta ctcaatagta tgactaatta aaaaaaaaaatt tttttctaa 600
aaaaaaaaaaaa aaaaaaa 616

Q337799 NAE 3377/99 PCT 20/59

<210> 10
<211> 1562
<212> genomic DNA
<213> Arabidopsis thaliana

<400> 10

ATGAAAAAAA	TATCTTCACA	TTATTGGTA	GTCATAGCGA	TACTCGTTGT	50	
GGTGACGATG	ACCTCGATGT	GTCAAGCTGT	GGGTAGCAAC	GTGTACCCCTT	100	
TGATTCTGGT	TCCAGGAAAC	GGAGGTAACC	AGCTAGAGGT	ACGGCTGGAC	150	
AGAGAATAACA	AGCCAAGTAG	TGTCTGGTGT	AGCAGCTGGT	TATATCCGAT	200	
TCATAAGAAG	AGTGGTGGAT	GGTTTAGGCT	ATGGTTCGAT	GCAGCAGTGT	250	
TATTGTCCTCC	CTTCACCAGG	TGCTTCAGCG	ATCGAATGAT	GTGTTACTAT	300	
GACCTGATT	TGGATGATTA	CCAAAATGCT	CCTGGTGTCC	AAACCCGGGT	350	
TCCTCATTTTC	GGTTCGACCA	AATCACTTCT	ATACACCGAC	CCTCGTCTCC	400	
GGTTAGTACT	TTCCAAGATA	TATCATTTCG	GGACACATTG	ATAATGAACA	450	
AAATAGACAT	AAATTGGGG	GATTATTGTT	ATATCAATAT	CCATTTATAT	500	
GCTAGTCGGT	AATGTGAGTG	TTATGTTAGT	ATAGCTTAATG	TGAGTGTAT	550	
GTGATTTTCC	ATTTAAATG	AAGCTAGAAA	GTTGTCGTTT	AATAATGTTG	600	
CTATGTCATG	AGAATTATAA	GGACACTATG	TAAATGTTAGC	TTAATAATAA	650	
GGTTTGATTT	GCAGAGATGC	CACATCTTAC	ATGGAACATT	TGCTGAAAGC	700	
TCTAGAGAAA	AAATGCCGGT	ATGTTAACGA	CCAAACCATC	CTAGGAGCTC	750	
CATATGATTT	CAGGTACGGC	CTGGCTGCTT	CGGGCCACCC	GTCGGTGTAA	800	
GCCTCACAGT	TCCTACAAGA	CCTCAAACAA	TTGGTGGAAA	AAACTAGCAG	850	
CGAGAACGAA	GGAAAGCCAG	TGATACTCCT	CTCCCATAGC	CTAGGAGGAC	900	
TTTTCTGCTCT	CCATTTCCTC	AACCCTTACCA	CCCCCTTCATG	GCGCCGCAAG	950	
TACATCAAAC	ACTTGTGTC	ACTCGCTGCG	CCATGGGTG	GGACGATCTC	1000	
TCAGATGAAG	ACATTGCTT	CTGGCAACAC	ACTCGGTGTC	CCTTTAGTTA	1050	
ACCCCTTGCT	GGTCAGACGG	CATCAGAGGA	CCTCCGAGAG	TAACCAATGG	1100	
CTACTTCCAT	CTACCAAAGT	GTTTCACCGAC	AGAACAAAC	CGCTTGTGCGT	1150	
AAACTCCCCAG	GTAAACTACA	CAGCTTACGA	GATGGATCGG	TTTTTTGCAG	1200	
ACATTGGATT	CTCACAAGGA	GTTGTCCTT	ACAAGACAAG	AGTGTGCGCT	1250	
TTAACAGAGG	AGCTGATGAC	TCCGGGAGTG	CCAGTCACCT	GCATATATGG	1300	
GAGAGGAGTT	GATACACCGG	AGGTTTGAT	GTATGGAAAA	GGAGGATTCG	1350	
ATAAGCAACC	AGAGATTAAG	TATGGAGATG	GAGATGGGAC	GGTTAATTG	1400	
GCGAGCTTAG	CAGCTTGA	AGTCGATAGC	TTGAACACCG	TAGAGATG	1450	
TGGAGTTTCG	CATACATCTA	TACTTAAAGA	CGAGATCGCA	CTTAAAGAGA	1500	
TTATGAAGCA	GATTCAATT	ATTAATTATG	AATTAGCCAA	TGTTAATGCC	1550	
GTCAATGAAT	GA					

1562

<210> 11
<211> 3896
<212> genomic DNA
<213> Arabidopsis thaliana

<400> 11	
ATGGGAGCGA ATTGCAAATC AGTAACGGCT TCCCTCACCG TCATGCCGT	50
TTTTTCTTG ATTTGCGGTG GCGGAACATGC GGTGGAGGAT GAGACCGAGT	100
TTCACGGCGA CTACTCGAAG CTATCGGGTA TAATCATTCC GGGATTGCG	150
TCGACGCAGC TACGAGCGTG GTCGATCCTT GACTGTCCAT ACACCTCCGT	200
GGACTTCAAT CCGCTCGACC TCGTATGGCT AGACACCACT AAGGTCCGTG	250
ATCTTCATTT CCTTCGCTCC TTATTCTGTC GGTCGAGTC CTTGTTGATG	300
AATTCCAAGC GAAATATAGC AATGAAGCAT GTCTCGTCTC TCTTATTGAT	350
TCGTTCATTA GTCAACAGTG ACGCTTCTGA ATCTGAGTT AGAGTCATAT	400
AAAACAGCTG ACTCGGCGAG TGTTTCCCAT CGCTTTGGT TCGCTAAATG	450
TAGCGCAATG AATGTGTAAT TAGTCTGCCG TTTTTATTCA ACTAGATCTG	500
CAAGTTTTC AGAGTGTCA ATAGTAGTTA GAAAATGTTA GGTCATTGTA	550
CTTGTGCATT GTGATTCTTT TGGTTGTTGC TTACTGATCG ACGTGATGGA	600
TGGTTTACAG CTTCTTCTG CTGTCAACTG CTGTTTAAG TGTATGGTC	650
TAGATCCTTA TAATCAAACA GACCATCCCG AGTGTAAAGTC ACGGCCTGAC	700
AGTGGTCTT CAGCCATCAC AGAATTGGAT CCAGGTTACA TAACAGGTAG	750
TTTCGGATTT TTCTTTCTTT TGAGTTTCTC TCAATTGAT ATCATCTGT	800
TGTGATATAA TATGGCTAAG TTCATTAAATT TGGTCAATT TCAGGTCCCTC	850
TTTCTACTGT CTGAAAGAG TGGCTTAAGT GGTGTGTTGA GTTTGGTATA	900
GAAGCAAATG CAATTGTCGC TGTTCCATAC GATTGGAGAT TGTCACCAAC	950
CAAATTGGAA GAGCGTGACC TTTACTTCA CAAGCTCAAG TTAGTCCTTA	1000
TCAGGCTAAT GTCTTTATC TTCTCTTTT ATGTAAGATA AGCTAAGAGC	1050
TCTGGTCGTC TTCTTTTG CAGGTTGACC TTTGAAACTG CTTTAAACT	1100
CCGTGGCGGC CCTTCTATAG TATTTGCCA TTCAATGGGT AATAATGTCT	1150
TCAGATACTT TCTGGAATGG CTGAGGCTAG AAATTGCACC AAAACATTAT	1200
TTGAAGTGGC TTGATCAGCA TATCCATGCT TATTTCGCTG TTGGTACCGG	1250
CCTACTATCC TTAAGTIAAC ATTTTATTT TTCTCTAATT GGGGGAGTTA	1300
TGTTGTGACT TACTGGATTG AGCTCGATAC CTGATTGTT GTTGATTGAT	1350
GAGCTCCTCT TCTGGTTCT CTTGAGGCAA TCAAATCTAC TCTCTGCGT	1400
GTAACGTTG GCCTTCCTGT TTCTGAGGTG ACCTCTGACT TCTCTTAGT	1450
TTTAAGTAGT TGATATCAAC CAGGTCTTAT AACTCACTGG ATTTTCTT	1500
TGAAAGTATT ACTTTGTTA ATTGAACATGC TGTACGCGAT ATGGTATCTG	1550
AGATCTTGA AGTGTAGTT ATCAAAGAAC ATATTGTGGG TAGTATACCT	1600
GTCAGCGGCC TTAGCTAATA CAACCAAAC ACATGTACAC TGATTAGTT	1650
TTCAAGATAT TATGGTAGAC TTAAAGTGA GAAGAAACTT TGACTGAAAT	1700
CTTTTTATT TAATAGGCTA TGATTGTT ATTGAAATCA TGTGACATAT	1750
TGACATGCGC TTCTCATGTT TTTTGTGGC AAGGCTTCAG GGAACCTGTC	1800
GGTTGTTGTC CAATTCTTT GCGTCGTAT TGTGGCTTAT GCCATTTC	1850
AAGAATTGCA AGGGTGATAA CACATTCTGG ACCATTTTT CTGGGGGTGC	1900
TGCAAAGAAA GATAAGCGCG TATACCACTG TGATGAAGAG GAATATCAAT	1950
CAAAATATTC TGGCTGGCCG ACAAAATATTA TTAACATTGA AATTCTTCC	2000
ACTAGCGGTT AGACTCTGTA TATGCAACTG TAACACTAAC AAAAGTTCA	2050
CCAAGAAATG TCACTCTCAT ATTCGTCTC TTTGATGTTGATCCATCACT	2100
TACAGAAACA GCTCTAGTCA ACATGACCG CATGGAATGT GGCCTTCCA	2150
CCCTTTGTC TTTCACAGCC CGTGAACCTAG CAGATGGGAC TCTTTCAAA	2200
GCAATAGAAG ACTATGACCC AGATAGCAAG AGGATGTTAC ACCAGTTAA	2250
GAAGTACGTA CCTTTCTTG TGATAAGAAA TATTGCTCAT CGATCATCAC	2300
TTGCTGGCTT CTTGTACGTC AAATTGTTT GTTAAATCT CTATATCAAT	2350
TGTTCATATG CTTTGTCTT CTTACTATAA GAAACAAGTA TAATCAGAAA	2400
CCTTATTATT GATTATCAGT TCTCTCCTTA TATTATGGAA TGTCTTTTC	2450
GTTTACAGTT ATGAATGCAA AAGGGGTAT TTTAGTTGAT TGATTCTCTC	2500
ATTCTCTAGT TTGTTTGCAC TAATAGCGTC AATTGTTTT TTCTAGCAAA	2550
TCTTTGTGAA TTATATATAA CATGCTAATC ATACTTTCA GGTTGTATCA	2600
TGATGACCC GTTTTAATC CTCTGACTCC TTGGGAGAGA CCACCTATAA	2650
AAAATGTATT TTGCATATAT GGTGCTCATC TAAAGACAGA GGTATGATGC	2700
ATTCTCAATA TCACATTATG CGTTGACTTT GTTATTATAT TCCCCATTG	2750
GTTTGCAATA TCTTTTGGAA TTATGATTAA TCTTCTCCCT TGCACTTAT	2800
GCTATTAAGC GTTAAAGGTA CTAAATGTAT GAAGCTGTCT GTCATAGGTT	2850

GGTTATTACT	TTGCCCAAG	TGGCAAACCT	TATCCTGATA	ATTGGATCAT	2900
CACGGATATC	ATTTATGAAA	CTGAAGGTC	CCTCGTGTCA	AGGTAATT	2950
CCGCAATGGC	AGAAGTAAAA	CAGGAAGGCA	AAGTCTCTG	TATCACTCTA	3000
GTGGCATGTT	ATCTCAGTTG	CATAAGCAA	TTATTAACAA	ACTAAAATT	3050
AAGTACTTTT	TTATCATTCC	TTTGAGCTT	AGTGGATGAT	CAGTGGCTTA	3100
AAGTGGGAAG	AGGTGTTGCA	TGAAACATGA	CACTGTATC	AAAGATAACT	3150
AGCAAAACAA	AACTAACCCA	TTTCTGAATT	TCATATTATT	AGGAGTAGTC	3200
GTGCTTTAA	AAAATTGTT	TTAACAAACC	GAAAAACTAG	TTCATATCTT	3250
GATTGTGCAA	TATCTGCAGG	TCTGGAACGT	TGGTTGATGG	GAACGCTGGA	3300
CCTATAACTG	GGGATGAGAC	GGTAAGCTCA	GAAGTTGGTT	TTGAAATTAT	3350
CTTCTTGCCTA	ACTACTGAAG	ACTAACATAA	TACTTGCTTC	TGGAACACTG	3400
CTTGCTATGT	TCTCTAGTAC	ACTGCAATAT	TGACTCTCCG	CTACTTTTAT	3450
TGATTATGAA	ATTGATCTCT	TATAGGTACC	CTATCATTCA	CTCTCTTGGT	3500
GCAAGAATTG	GCTCGGACCT	AAAGTTAACAA	TAACAAATGGC	TCCCCAGGTA	3550
CTCTTTTTA	GTTCTCACC	TTATATAGAT	CAAACTTTAA	GTTGACTTTT	3600
CTGGTTATGT	GTTGATTTAC	CTCCAATTG	TTCTTTCTAA	AAATCATATA	3650
TCTCTGTACT	CCTCAAGAAC	TTGTATTAAT	CTAAACGAGA	TTCTCATGG	3700
GAAAATAAAA	CAACAGCCAG	AACACGATGG	AAGCGACGTA	CATGTGGAAC	3750
TAAATGTTGA	TCATGAGCAT	GGGTCAGACA	TCATAGCTAA	CATGACAAAA	3800
GCACCAAGGG	TTAAGTACAT	AACCTTTAT	GAAGACTCTG	AGAGCATTC	3850
GGGAAAGAGA	ACCGCAGTCT	GGGAGCTTGA	TAAAAGTGGG	TATTAA	3896

Sequence Data from ABI Prism 377 DNA Sequencer

<210> 12
<211> 709
<212> cDNA
<213> Lycopersicon esculentum

<400> 12

CTGGGGCCAA AAGTGAACAT AACAGGACA CCACAGTCAG AGCATGATGT	50
TCAGATGTAC AAGTGCATCT AAATATAGAG CATCAACATG GTGAAGATAT	100
CATTCCCAAT ATGACAAAGT TACCTACAAT GAAGTACATA ACCTATTATG	150
AGGATTCTGA AAGTTTCCA GGGACAAGAA CAGCAGTTG GGAGCTTGAT	200
AAAGCAAATC ACAGGAACAT TGTCAGATCT CCAGCTTGA TGCAGGGAGCT	250
GTGGCTTGAG ATGTGGCATG ATATTCATCC TGATAAAAAG TCCAAGTTTG	300
TTACAAAAGG TGGTGTCTGA TCCTCACTAT TTTCTTCTAT AAATGTTGTA	350
GTTTGTATTG ACATTTGTAAG TATTGCAACA AAAAGCAAAG CGTGGGCCTC	400
TGAGGGATGA GGACTGCTAT TGGGATTACG GGAAAGCTCG ATGTGCATGG	450
GCTGAACATT GTGAATACAG GTTAGAATAT TCAAATTATA TTTTGCAAAA	500
TATTCTCTTT TTGTTGATTT AGGCCACCTT TCCCCGGTCA CAACGATGCA	550
GATATGTATT CGGGGATGTT CACCTGGGAC AGAGTTGCAG ATTGAAGAGT	600
TCTACATCTC ACATCCTGTC ACACATATGTG TGATATTAA GAAACTTTGT	650
TGCGCGAAC ACAAGTTG CACAAACATT TGAAGAAGAA AGCGAAATGA	700

TTCAGAGAG

709

<210> 13
<211> 623
<212> PRT
<213> *Schizosaccharomyces pombe*

<400> 13

MASSKKSKTHKKKKEVKSPIDLPSKKPTRALSEQPSASETQSVSNKSRKSFGKRLNFILGAILGICGA 70
FFFAGDDNAVFDPATLDKFGNMLGSSDLFDDDIKGYLISYNVFKDAPFTTDKPSQSPSGNEVQVGLDMYNE 140
GYRSDHPVIMVPGVISSGLEWSFNNCSPYFRKRLWGSWSMLKAMFLDKQCWLHMLDKKTGLDPKGI 210
KLRAAQGFEAADFFITGYWIWSKVIEENLAAIGYEPNNMLSASYDWRLSYANLEERDKYFSKLKMFIEYSN 280
IVHKKKVVVLISHSMGSQVTYYFFKWVEAEGYGNGGPTWVNDHIEAFINISGLIGAPKTVALLSGEMKD 350
TGIVITLNILEKFFSRSERAMMVRTMGGVSSMLPKGGDVAPDDLQNTNFSNGAIIRYREDIDKDHDEFDI 420
DDALQFLKNVTDDDFKVMLAKNYSHGLAWTEKEVLKNNEMPSKWINPLETSLPYAPDMKIYCVHGVGKPT 490
ERGYYYTNNEGQPVIDSSVNDGTKVENGIVMDDGDTLPILALGLVCNKVWQTKRFPANTSITNYEIK 560
HEPAAFDLRGGRSAEHVDILGHSELNEIILKVSSGHGDSVPNRYISDIQEIIINEINLDKPRN 623

<210> 14
<211> 432
<212> PRT
<213> *Arabidopsis thaliana*

<400> 14
MKKISSHYSVVIAILVVVTMTSMCQAVGSNVPLILVPGNGGNQLEVRLDREYKPSSWCSSWLYPIHKK 70
SGGWFRWLWFDAAVLLSPFTRCFSDRMMLYYDPDLDDYQNAPGVQTRVPHFGSTKSLLYLDPRLRDATSYM 140
EHLVKALEKKCGYVNDQTILGAPYDFRYGLAASGHPSRVASQFLQDLKQLVEKTSSNEGKPVILLSHSL 210
GGLFVLFHFLNRTTPSWRRKYIKHFVALAAPWGKTISQMKTFAASGNTLGVPLVNPLLVRHQRTSESNQWL 280
LPSTKVFHDRTKPLVVTPQVNNTAYEMDRFFADIGFSQGVVPYKTRVLPLTEELMTPGVPTCIYGRGVD 350
TPEVLMYGGFDKQPEIKYGDGDGTVNLASLAALKVDSLNTVEIDGVSHTSILKDEIALKEIMKQISII 420
NYELANVNAVNE

432

<210> 15

<211> 552

<212> PRT

<213> *Arabidopsis thaliana*

<400> 15

MGANSKSVTASFTVIAVFFLICGGRTAVEDETEFHGDYSKLSGIIIPGFASTQLRAWSILD
PLDLVWLDTPTKLLSAVNCFKCMVLDPYNQTDHPECKSRPDGLSAITEADPGYITGPLSTV
WKEWLKWC 140
VEFGIEANAIIVAVPYDWRLSPTKLEERDLYFHKLKLTFETALKLRGGPSIVFAHSMGN
NVFRYFLEWLRL 210
EIAPKHYLKWLQHIIAYFAVGAPLLGSVEAIKSTLSGVTFGLPVSEGTA
RLLSNSFASSLWLMPFSKNC 280
KGDNTRFWTHFSGGAAKKDKRVYHCDEEEYQSKYSGWPTNIINIEIPSTSARELADGT
LFKAIEDYDPDSK 350
RMLHQLKKYVPFFVIRNIAHRSSLAGFLYHDDPVFNPLTPWERPPIKNVFCIYGA
HLKTEVGYYFAPSG 420
KPYPDNWIITDIYETEGSLVSRSGTVVDGNAGPITGDETVPYHSLSWCKNWLGPKVN
ITMAPQILIGKI 490
KQQPEHDGSDVHVELNVDHEHSDIIANMTKAPRVKYITFYEDSESIPGKRTAVWE
LDKSGY 552

5

10

15

<170> PatentIn Ver. 2.0

<210> 1Q
<211> 66120 <212> PRT
<213> *Saccharomyces cerevisiae*

<400> 1

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
25 1 5 10 15Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
20 25 3030 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
35 40 45Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
50 55 6035 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
65 70 75 8040 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
85 90 95Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
100 105 11045 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
115 120 125Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
130 135 14050 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly
145 150 155 16055 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
165 170 175Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile
180 185 19060 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
195 200 205

Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
 210 215 220

5 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
 225 230 235 240

Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
 245 250 255

10 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
 260 265 270

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
 15 275 280 285

Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
 290 295 300

20 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
 305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
 25 325 330 335

Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Arg Gly Trp Val
 340 345 350

Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
 30 355 360 365

Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
 370 375 380

Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
 35 385 390 395 400

Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
 40 405 410 415

Met Leu Pro Lys Gly Glu Val Ile Trp Gly Asp Met Lys Ser Ser
 420 425 430

Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
 45 435 440 445

Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
 450 455 460

Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
 50 465 470 475 480

Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
 55 485 490 495

Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
 500 505 510

Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
 60 515 520 525

Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
 530 535 540

Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
 5 545 550 555 560

Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
 565 570 575

10 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
 580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
 595 600 605

15 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
 610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
 20 625 630 635 640

Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
 645 650 655

25 Met Pro Phe Pro Met
 660

30 <210> 20
 <211> 387
 <212> PRT
 <213> Arabidopsis thaliana

<400> 2
 35 Val Gly Ser Asn Val Tyr Pro Leu Ile Leu Val Pro Gly Asn Gly Gly
 1 5 10 15

Asn Gln Leu Glu Val Arg Leu Asp Arg Glu Tyr Lys Pro Ser Ser Val
 40 20 25 30

Trp Cys Ser Ser Trp Leu Tyr Pro Ile His Lys Lys Ser Gly Gly Trp
 35 40 45

Phe Arg Leu Trp Phe Asp Ala Ala Val Leu Leu Ser Pro Phe Thr Arg
 45 50 55 60

Cys Phe Ser Asp Arg Met Met Leu Tyr Tyr Asp Pro Asp Leu Asp Asp
 65 70 75 80

50 Tyr Gln Asn Ala Pro Gly Val Gln Thr Arg Val Pro His Phe Gly Ser
 85 90 95

Thr Lys Ser Leu Leu Tyr Leu Asp Pro Arg Leu Arg Asp Ala Thr Ser
 55 100 105 110

Tyr Met Glu His Leu Val Lys Ala Leu Glu Lys Lys Cys Gly Tyr Val
 115 120 125

60 Asn Asp Gln Thr Ile Leu Gly Ala Pro Tyr Asp Phe Arg Tyr Gly Leu
 130 135 140

Ala Ala Ser Gly His Pro Ser Arg Val Ala Ser Gln Phe Leu Gln Asp
 145 150 155 160

5 Leu Lys Gln Leu Val Glu Lys Thr Ser Ser Glu Asn Glu Gly Lys Pro
 165 170 175

Val Ile Leu Leu Ser His Ser Leu Gly Gly Leu Phe Val Leu His Phe
 180 185 190

10 Leu Asn Arg Thr Thr Pro Ser Trp Arg Arg Lys Tyr Ile Lys His Phe
 195 200 205

Val Ala Leu Ala Ala Pro Trp Gly Gly Thr Ile Ser Gln Met Lys Thr
 210 215 220

15 Phe Ala Ser Gly Asn Thr Leu Gly Val Pro Leu Val Asn Pro Leu Leu
 225 230 235 240

20 Val Arg Arg His Gln Arg Thr Ser Glu Ser Asn Gln Trp Leu Leu Pro
 245 250 255

Ser Thr Lys Val Phe His Asp Arg Thr Lys Pro Leu Val Val Thr Pro
 260 265 270

25 Gln Val Asn Tyr Thr Ala Tyr Glu Met Asp Arg Phe Phe Ala Asp Ile
 275 280 285

Gly Phe Ser Gln Gly Val Val Pro Tyr Lys Thr Arg Val Leu Pro Leu
 290 295 300

30 Thr Glu Glu Leu Met Thr Pro Gly Val Pro Val Thr Cys Ile Tyr Gly
 305 310 315 320

Arg Gly Val Asp Thr Pro Glu Val Leu Met Tyr Gly Lys Gly Phe
 325 330 335

Asp Lys Gln Pro Glu Ile Lys Tyr Gly Asp Gly Asp Gly Thr Val Asn
 340 345 350

40 Leu Ala Ser Leu Ala Ala Leu Lys Val Asp Ser Leu Asn Thr Val Glu
 355 360 365

Ile Asp Gly Val Ser His Thr Ser Ile Leu Lys Asp Glu Ile Ala Leu
 370 375 380

45 Lys Glu Ile
 385

50 <210> 30
 <211> 389
 <212> PRT
 <213> Arabidopsis thaliana

55 <400> 3
 Leu Lys Lys Glu Gly Leu Lys Ala Lys His Pro Val Val Phe Ile Pro
 1 5 10 15

60 Gly Ile Val Thr Gly Gly Leu Glu Leu Trp Glu Gly Lys Gln Cys Ala
 20 25 30

	Asp	Gly	Leu	Phe	Arg	Lys	Arg	Leu	Trp	Gly	Gly	Thr	Phe	Leu	Cys	Trp
			35					40					45			
5	Val	Glu	His	Met	Ser	Leu	Asp	Asn	Glu	Thr	Gly	Leu	Asp	Pro	Ala	Gly
		50				55						60				
	Ile	Arg	Val	Arg	Ala	Val	Ser	Gly	Leu	Val	Ala	Ala	Asp	Tyr	Phe	Ala
		65				70				75			80			
10	Pro	Gly	Tyr	Phe	Val	Trp	Ala	Val	Leu	Ile	Ala	Asn	Leu	Ala	His	Ile
			85					90				95				
	Gly	Tyr	Glu	Glu	Lys	Asn	Met	Tyr	Met	Ala	Ala	Tyr	Asp	Trp	Arg	Leu
		100					105					110				
15	Ser	Phe	Gln	Asn	Thr	Glu	Arg	Asp	Gln	Thr	Leu	Ser	Arg	Met	Lys	Ser
		115					120					125				
20	Asn	Ile	Glu	Leu	Met	Val	Ser	Thr	Asn	Gly	Gly	Lys	Lys	Ala	Val	Ile
		130				135					140					
	Val	Pro	His	Ser	Met	Gly	Val	Leu	Tyr	Phe	Leu	His	Phe	Met	Lys	Trp
		145				150				155			160			
25	Val	Glu	Ala	Pro	Ala	Pro	Leu	Gly	Gly	Gly	Gly	Pro	Asp	Trp	Cys	
		165					170					175				
	Ala	Lys	Tyr	Ile	Lys	Ala	Val	Met	Asn	Ile	Gly	Gly	Pro	Phe	Leu	Gly
		180					185					190				
30	Val	Pro	Lys	Ala	Val	Ala	Gly	Leu	Phe	Ser	Ala	Glu	Ala	Lys	Asp	Met
		195					200					205				
	Arg	Met	Thr	Arg	Thr	Trp	Asp	Ser	Thr	Met	Ser	Met	Leu	Pro	Lys	Gly
		210				215				220						
	Gly	Asp	Thr	Ile	Trp	Gly	Gly	Leu	Asp	Trp	Ser	Pro	Glu	Leu	Pro	Asn
		225				230				235			240			
40	Ala	Pro	Glu	Met	Glu	Ile	Tyr	Ser	Leu	Tyr	Gly	Val	Gly	Ile	Pro	Thr
			245				250					255				
	Glu	Arg	Ala	Tyr	Val	Tyr	Lys	Leu	Asn	Gln	Ser	Pro	Asp	Ser	Cys	Ile
		260				265						270				
45	Pro	Phe	Gln	Ile	Phe	Thr	Ser	Ala	His	Glu	Glu	Asp	Glu	Asp	Ser	Cys
		275				280						285				
50	Leu	Lys	Ala	Gly	Val	Tyr	Asn	Val	Asp	Gly	Asp	Glu	Thr	Val	Pro	Val
		290				295					300					
	Leu	Ser	Ala	Gly	Tyr	Met	Cys	Ala	Lys	Ala	Trp	Arg	Gly	Lys	Thr	Arg
		305				310				315			320			
55	Phe	Asn	Pro	Ser	Gly	Ile	Lys	Thr	Tyr	Ile	Arg	Glu	Tyr	Asn	His	Ser
			325					330					335			
	Pro	Pro	Ala	Asn	Leu	Leu	Glu	Gly	Arg	Gly	Thr	Gln	Ser	Gly	Ala	His
		340				345						350				
60	Val	Asp	Ile	Met	Gly	Asn	Phe	Ala	Leu	Ile	Glu	Asp	Ile	Met	Arg	Val

	355	360	365		
				Ala Ala Gly Gly Asn Gly Ser Asp Ile Gly His Asp Gln Val His Ser	
	370	375	380		
5	Gly Ile Phe Glu Trp				
	385				
10	<210> 4A <211> 1986 <212> DNA <213> <i>Saccharomyces cerevisiae</i>				
15	<220> <221> CDS <222> (1)...(1983)				
20	<400> 4 atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser 1 5 10 15				48
25	gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg 20 25 30				96
30	aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly 35 40 45				144
35	att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg 50 55 60				192
40	aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu 65 70 75 80				240
45	att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe 85 90 95				288
50	ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe 100 105 110				336
55	gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val 115 120 125				384
60	ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn 130 135 140				432
	tac tcc aca tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly 145 150 155 160				480
	aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val				528

	165	170	175	
5	atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile 180 185 190			576
10	gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp 195 200 205			624
15	gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp 210 215 220			672
20	ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn 225 230 235 240			720
25	ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile 245 250 255			768
30	gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile 260 265 270			816
35	ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu 275 280 285			864
40	gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys 290 295 300			912
45	gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta Glu Gln Ile Glu Leu His Gln Leu Ser Gly Glu Lys Val Cys Leu 305 310 315 320			960
50	att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp 325 330 335			1008
55	gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Arg Gly Trp Val 340 345 350			1056
60	aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly 355 360 365			1104
	gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr 370 375 380			1152
	att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser 385 390 395 400			1200
	aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser 405 410 415			1248

	atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser 420 425 430	1296
5	tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile 435 440 445	1344
10	cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met 450 455 460	1392
15	aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln 465 470 475 480	1440
20	aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu 485 490 495	1488
25	gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met 500 505 510	1536
30	gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr 515 520 525	1584
35	ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp 530 535 540	1632
40	gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro 545 550 555 560	1680
45	gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser 565 570 575	1728
50	atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga Met Cys His Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly 580 585 590	1776
55	att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp 595 600 605	1824
60	ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser 610 615 620	1872
65	gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp 625 630 635 640	1920
70	ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln 645 650 655	1968

1986

atg ccc ttc cca atg taa			
Met Pro Phe Pro Met			
660			
5			
<210> 5 α			
<211> 661			
<212> PRT			
<213> <i>Saccharomyces cerevisiae</i>			
10			
<400> 5			
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser	1	5	10
			15
15			
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg	20	25	30
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly	35	40	45
20			
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg	50	55	60
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu	65	70	75
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe	85	90	95
25			
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe	100	105	110
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val	115	120	125
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn	130	135	140
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly	145	150	155
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val	165	170	175
30			
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile	180	185	190
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp	195	200	205
35			
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp	210	215	220
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn	225	230	235
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile	245	250	255
40			
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile	260	265	270

Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
 275 280 285
 5 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
 290 295 300
 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
 305 310 315 320
 10 Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
 325 330 335
 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
 15 340 345 350
 Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
 355 360 365
 20 Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
 370 375 380
 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
 385 390 395 400
 25 Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
 405 410 415
 Met Leu Pro Lys Gly Glu Val Ile Trp Gly Asp Met Lys Ser Ser
 30 420 425 430
 Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
 435 440 445
 35 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
 450 455 460
 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
 465 470 475 480
 40 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
 485 490 495
 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
 45 500 505 510
 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
 515 520 525
 50 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
 530 535 540
 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
 545 550 555 560
 55 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
 565 570 575
 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
 60 580 585 590

Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
595 600 605

5 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
610 615 620

Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
625 630 635 640

10 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
645 650 655

Met Pro Phe Pro Met
660

15

SEQUENCE LISTING

<110> Scymnus Dr., Sten
<120>

<130>
<140>
<141>
<160>

<210> 1^b

<211> 1986

<212> genomic DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> CDS

<222> (1)..(1983)

<400> 1

atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct	48
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser	
1 5 10 15	

gat gan aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga	96
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg	
20 25 30	

aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt	144
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly	
35 40 45	

att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg	192
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg	
50 55 60	

aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg	240
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu	
65 70 75 80	

att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt	288
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe	
85 90 95	

ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt	336
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe	
100 105 110	

gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt		384	
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val			
115	120	125	
ctc cca caa ggt ata agt tcg ttt acc gat gat att cag gct ggt aac		432	
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn			
130	135	140	
cac tcc aca tct tcc tta gat gat ctc agt gaa aat ttt gcc gtt ggt		480	
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly			
145	150	155	160
aaa caa ctc tta cgt gat tat aat acc gag gcc aaa cat ccc gll yea		520	
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val			
165	170	175	
atg gct cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att		576	
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile			
180	185	190	
gga gag gat gag tgc gat agt tct ggg cat ttt cgt aaa ccc ctg tgg		624	
Gly Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp			
195	200	205	
gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tct tgg		672	
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp			
210	215	220	
ctg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac		720	
Ieu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn			
225	230	235	240
ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc		768	
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile			
245	250	255	
gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att		816	
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile			
260	265	270	
ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt		864	
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu			
275	280	285	

GCGTGTCTTGTTGTCGCG

gca tac tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys 290 295 300	912
gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu 305 310 315 320	950
att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg Ile Gly His Ser Met Gly Ser Gln Ile Phe Tyr Phe Met Lys Trp 325 330 335	1006
gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg stt Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Arg Gly Trp Val 340 345 350	1056
aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly 355 360 365	1104
gct cca aag gca gtt cca gct cta att agt ggt gaa atg aza gat acc Ala Pro Lys Ala Val Pro Ala Leu Ser Gly Glu Met Lys Asp Thr 370 375 380	1152
att caa tta aat acg tca gcc atg tat ggt ttg gaa aag ttc ttc tca Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser 385 390 395 400	1200
aga att gag aga gta aza atg tta caa acg tgg ggt ggt atz cca tca Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser 405 410 415	1248
atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct Met Leu Pro Lys Gly Glu Val Ile Trp Gly Asp Met Lys Ser Ser 420 425 430	1296
tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile 435 440 445	1344
cga ttt gaa agg aat acg aac gat gct ttc aac aza aat ttg aca atg Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met 450 455 460	1392
aaa gac gcc att aac atg aca tta tcg atz tca cct gaa tgg ctc caa Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln 465 470 475 480	1440

aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa		1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu		
485	490	495
gag cta aga ana aat gag cta cac cac aag cac tgg tcg aat cca atg		1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met		
500	505	510
gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac		1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr		
515	520	525
ggg stg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat		1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp		
530	535	540
gac tcc tct gct ctg aat ttg acc acc gac tac gaa agc aag caa cct		1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro		
545	550	555
gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca		1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser		
565	570	575
atg tgt cac aaa cgg gcc cag ggt gct tca ccg tac aac cct gcc gga		1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly		
580	585	590
att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat		1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp		
595	600	605
ata cgt ggc gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc		1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser		
610	615	620
scg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat		1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp		
625	630	635
ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag		1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln		
645	650	655

1986

atg ccc ttc cca atg taa
Met Pro Phe Pro Met
660

<210> 2f9

<211> 661

<212> FRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
 1 5 10 15
 Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
 20 25 30
 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
 35 40 45
 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
 50 55 60
 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
 65 70 75 80
 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
 85 90 95
 Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
 100 105 110
 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
 115 120 125
 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
 130 135 140
 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly
 145 150 155 160
 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
 165 170 175
 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile
 180 185 190
 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
 195 200 205
 Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
 210 215 220
 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
 225 230 235 240
 Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
 245 250 255
 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
 260 265 270
 Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
 275 280 285
 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
 290 295 300
 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
 305 310 315 320

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
 325 330 335
 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
 340 345 350
 Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
 355 360 365
 Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
 370 375 380
 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
 385 390 395 400
 Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
 405 410 415
 Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
 420 425 430
 Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
 435 440 445
 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
 450 455 460
 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
 465 470 475 480
 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
 485 490 495
 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
 500 505 510
 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
 530 535 540
 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
 545 550 555 560
 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
 565 570 575
 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
 580 585 590
 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
 595 600 605
 Ile Arg Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
 610 615 620
 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
 625 630 635 640
 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
 645 650 655
 Met Pro Phe Pro Met
 660

<210> 3
<211> 2312

<212> genomic DNA

<213> Schizosaccharomyces pombe

<400> 3

ATGGCGTCTT CCAAGAACAG CAAARCTCAT AAGAAAAAGA AAGAAGTCAA	50
ATCTCCATAC GACTTACCAA ATTCAAAGAA ACCAACTCGC GCTTTGAGTG	100
AGCAACCTTC AGCGTCCGAA ACACAATCTG TTCAAATAA ATCAAGAAAA	150
TCTAAATTG GAAAAGATT GAATTTATA TTGGGCGCTA TTTGGGAAT	200
ATCCGGTGCT TTTTTTTTCG CTGTTGGAGA CGACAATGCT GTTTTCGACC	250
CTGCTACGTT AGATAAATTG GGGAAATATGC TAGGCTCTTC AGACTTGT	300
GATCACATTA AACGATATTT ATCTTATAAT GTGTTAAGG ATGCACCTTT	350
TACTACGGAC AAGCCTTCGC ACTCTCCTAG CGGAAATGAA GTTCAAGTTG	400
GCTTGATAT GTACAATGAG GGATATCGAA GTGACCATCC TGTATTATG	450
GTTCCTGGTG TTATCAGCTC AGGATAGAA AGTTGGTCGT TTAATAATTG	500
CTCGATTCCCT TACTTTAGGA AACGCTTTG GGGTAGCTGG TCTATGCTGA	550
AGGCAATGTT CCTTGACAG CAATGCTGGC TTGAACATTT AATGCTTGAT	600
AAAAAAACCG GCTTGGATCC GAAGGAAATT AAGCTGCGAG CAGCTCAGGG	650
GTGGAAAGCA GCTGATTTT TTATCACGGG CTATTGGATT TGGAGTAAAG	700
TAATTGAAAA CCTTGCTGCA ATTGGTTATG AGCCTAATAA CATGTTAAGT	750
GCTTCTTACG ATTGGCGGTT ATCATATGCA AATTTAGAGG AACGTGATAA	800
ATATTTTCA AAGTTAAAAA TGTTCATTGA GTACAGCAAC ATTGTACATA	850
AGAABAAACGT AGTGGTGTATT TCTCACTCCA TGGGTTACACA GGTTACGTAC	900
TATTTTTTA AGTGGGTTGA AGCTGAGGGC TACGGAAATG GTGGACCGAC	950
TTGGGTTAAT GATCATATTG AAGCATTATG AAATGTGAGT CTCGATGGTT	1000
GTGGACTAC GTTTCTAATCT TITGAATAGA TATCGGGATC TTTGATTGGA	1050
GCACCCAAAA CAGTGGCAGC GCTTTATCG GGTGAAATGA AAGATACAGG	1100
TATIGTAATT ACATTAACAA TGTTAATATT TAATTTTGC TAACCGTTT	1150
AAGCTCAATT GAATCACTTT TCGGTCTATG GGTAAAGCAAT AAATTGTTGA	1200
GATTTGTTAC TAATTTACTG TTAGTTTGG AAAAATTGTT TTCCCCCTCT	1250
GAGGTATATT CAAAAATACA AATGTGCTCT ACTTTTCTA ACTTTAATA	1300
GAGAGCCATG ATGGTTCGCA CTATGGGAGG AGTTAGTTCT ATGCTTCCTA	1350
AAGGAGGCCGA TGTGTATGG GGAATGCCA GTTGGTAAG AAATAATGTGC	1400

GCGGCCGC-GTCCTCCGCG

TGTTAATTTT TTATTAATAT TTAGGCTCCA GATGATCTTA ATCAAACAAA	1450
TTTTTCCAAT GGTGCAATTAA TTCGATATAG AGAAGAGACATT GATAAGGACC	1500
ACGATGAATT TGACATAGAT GATGCATTAC AATTTTTAAA AAATGTTACA	1550
GATGACCGATT TTAAGTCAT CCTAGCGAAA AATTATTCCC ACGGTCTTGC	1600
TTGGACTGAA AAGGAAGTGT TAAAAAATAA CGAAATGCCG TCTAAATGGA	1650
TTAATCCGCT AGAAGTAAGA ACATTAAGT TACTAAATTA TACTAACCCA	1700
ATAAGACTAG TCTTCCTTAT GCTCCTCGATA TGAAAATTAA TTGCGTTCAC	1750
CGGGTCCGAA AACCAACTGA GAGAGGTTAT TATTATACTA ATAATCCTGA	1800
GGGGCAACCT GTCATTGATT CCTCGGTTAA TGATGGAACA AAAGTGAAA	1850
ATGTGAGAGA ATTATGTTT CAAACATTCT ATTAACTGTT TTATTAGGTT	1900
AATGTTATGG ATGATGGTGA TGGAACTTTA CCAATATTAG CCCTGGTTT	1950
GGTGTCCAAT AAAGTTGGC AAACAAAAAG GTTTAACCT GCTAAATACAA	2000
GTATCACAAA TTATGAAATC AAGCATGAAC CTGCTGCGTT TGATCTGAGA	2050
GGAGGACCTC GCTCGGCAGA ACACGTCGAT ATACTTGGAC ATTCAAGAGCT	2100
AATGTATGT TCATTITACC TTACAAATTCT CTATTACTAA CTCTTGAAT	2150
AAGGAATTA TTTAAAGT TTCATCAGGC CATGGTGACT CGGTACCAAA	2200
CCGTTATATA TCAGATATCC AGTAAGGACA TAAAGTTTGT AGATTCGAAT	2250
TMACTAACTA ACCGAACAGG GAATAATAA ATGAGATAAA TCTCGATAAA	2300
CCTAGAAATT AA	2312

<210> 4b
 <211> 3685
 <212> genomic DNA
 <213> Arabidopsis thaliana
 <400> 4

ATGCCCTTA	TTCATCGGAA	AAACCCGACG	GAGAAACCAT	CGACGCCGCC	50
ATCTGAACAG	GTGGTGCACG	ATGAGGATTG	GCAAAAGAAA	CCACACCAAT	100
CTTCCAAATC	CCACCATAAG	AAATCGAACG	GAGGAGGGAA	GTGGTCGTGC	150
ATCGATTCTT	GTTCCTGGTT	CATTGGGTGT	GTGTGTGTAA	CCTGGTGGTT	200
TCTTCTCTTC	CTTTACAACG	CAATGCCCTGC	GAGCTTCCT	CAGTATGTAA	250
CGGAGCGAAT	CACGGGTCT	TTGCCTGACC	CGCCCGGTGT	TAAGCTCAAA	300
AAAGAAGGTC	TTAAGGGCAA	ACATCCTGTT	GTCTTCATTC	CTGGGATTGT	350
CACCGCTGGG	CTCGAGCTTT	GGGAAGGCAA	ACAATGCGCT	GATGGTTTAT	400
TTAGAAAACG	TTTGTGGGT	GGAACTTTG	GTGAAAGTCTA	CAAAAGGTGA	450
GCTCAACAAT	TCTCACTCTT	CCTTTATATT	GGGATTTGGA	TTGGATCTGA	500
TGAGATCACG	CACTTGTGTC	TTCTTCACA	TCACTCAAAAC	TTTAATTCCA	550
TGTTTGTCTG	TCTTACTCTT	TACTTTTTT	TTTTTTGAT	GTGAAACGCT	600
ATTTTCTTAA	GAGACTATTT	CTGTATGTGT	AAGGTAAGCG	TTCCAAGGAC	650
GTAATTGGCT	TGGACTATTT	CTGTTTGATT	GTAAACTTTA	GGATATAAAA	700
TAGCTGCCCT	GGAATTCAA	GTCATCTTAT	TGCCAAATCT	GTTGCTAGAC	750
ATGCCCTAGA	GTCCGTTCAT	AAACAAGTTAC	TTCCCTTACT	GTGGTGGGT	800
GTAGATTTAG	CTTGTGTAG	CGTATAATGA	AGTAGTGT	TATGTTTGT	850
TGGGAATAGA	GAAGTTCTAA	CTACATCTGT	GGAAAGTGTG	TTCAAGGCTGT	900
GATAGAGGAC	TGTTGCTTTA	TTATTCAACT	ATGTATATGT	GTAATTAAAG	950
CTAGTTCCCT	TTTGATCTTT	CAGCTCAATG	TGCTTTCTC	AATTTTTTC	1000
TCAATTCAA	AGTTTCACAT	CGAGTTTATT	CACATGTCTT	GAATTTCGTC	1050
CATCCTCGTT	CTGTTATCCA	GCTTTGAACT	CCTCCCCGACC	CTGCTATGGA	1100
TATATTAAAA	AAAAAGTGT	TTGTGGGTG	CATCTTGTT	ACGATCTGCA	1150
TCTCTTCCTT	TCGGCTCAGT	GTTCATGTTT	TTGCTATGGT	AGAGATGGGC	1200
ATGTTTATG	TTGATGGTAA	CAGGGTATA	GTTGATAGTA	TCTTAACCAA	1250
TCAATTATCT	CTTIGAITCA	GGCCTCTATG	TTGGGTGGAA	CACATGTCAC	1300
TTGACAATGA	AACTGGGTG	GATCCAGCTG	GTATTAGAGT	TCGAGCTGTA	1350

TCAGGACTCG TGGCTGCTGA CTACTTTGCT CCTGGCTACT TTGTCTGGGC 1400
 ACTGCTGATT CCTAACCTTG CACATATTGG ATATGAAGAG AAAAATAATGT 1450
 ACATGGCTGC ATATGACTGG CGGCTTCGT TTCAGAACAC AGAGGTTCTT 1500
 TTCTCATCGT TCTTCTTATT ATTCTGTTCC ATGTTACGTT TCTTCTTCA 1550
 TTACTTAAGG CTTAAATAATG TTTCATGTTG AATTAATAGG TACGTGATCA 1600
 GACTCTTAGC CGTATGAAAA GTAATATAGA GTTGATGGTT TCTACCAAACG 1650
 GTGGAAAAAAA AGCAGTTATA GTTCCGCATT CCATGGGGGT CTTGTATTIT 1700
 CTACATTTA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGGGGGGG 1750
 TGGGCCAGAT TGGTGTGCAA AGTATATTAA GCGGGTGTG AACATTGGTG 1800
 GACCATTCT TGGTGTGCTCA AAAGCTGTTG CAGGGCTTTT CTCTGCTGAA 1850
 GCAGGATG TTGCAGTTGC CAGGTATTGA ATATCTGCTT ATACTTTGAA 1900
 TGATCACAAAC CTTCGCTCTG GAACTCAAAG TTATTCTACT AAATATCAAT 1950
 TCTAATAACA TTGCTATATT ATCGCTGCAA CTGACATTGG TTGATTATIT 2000
 TTGCTGCTTA TGTAACTGAA ACTCTCTTGA GATTAGACAA ATGATGAATT 2050
 GATAATTCTT ACGCATTGCT CTGTGATGAC CAGTTCTTA GCTTCGACGA 2100
 TAACATTGTT CATACTGTCT TTTGGAGGGC ATTGAATTTC GCTATGGAAA 2150
 GCGCTGGAGC TTCCATGCTT GCATTCTTA CCAATTAGCG TTATTCTGCT 2200
 TCTTCAATT TTCTTGTATA TGCATCTATG GTCTTTATT TCTTCTTAAT 2250
 TAAAGACTCG TTGGATTAGT TGCTCTTAA GTCACTTGGT TCCTTAATAT 2300
 AGAAGTTAC TTTCTTCGAA AATTGCAGAG CGATTGCCCC AGGATTCTTA 2350
 GACACCGATA TATTTAGACT TCAGACCTTG CAGCATGTAA TGAGAATGAC 2400
 ACGCACATGG GACTCAACAA TGTCTATGTT ACCGAAGGGA GGTGACACGA 2450
 TATGGGGCGG CCTTGATTGG TCACCGGAGA AAGGCCACAC CTGTTGTGGG 2500
 AAAAGCRAA AGAACACCGA AACTTGTGGT GAAGCAGGTG AAAACGGAGT 2550
 TTCCAAGAAA AGTCCTGTTA ACTATGGAAG GATGATATCT TTTGGAAAG 2600
 AAGTAGCAGA CGCTGCGCCA TCTGAGATTA ATAATATTGA TTTTCGAGTA 2650
 AGGACATATA AATCATAATA AACCTTGTAC ATTTTGTGAT TGTATGATGA 2700
 ATATCTGTAC ATTTTATCTG GTGAAGGGTG CTGTCAAAGG TCAGAGTATC 2750
 CCAANTCACA CCTGTCGTGA CGTGTGGACA GAGTACCATG ACATGGGAAT 2800
 TGCTGGGATC AAAGCTATCG CTGAGTATAA GGTCTACACT GCTGGTGAAG 2850

3000 2500 2000 1500 1000

CTATAGATCT ACTACATTAT GTTGCTCCTA AGATGATGGC GCGTGGTGCC	2900
GCTCATTCT CTTATGGAA TGCTGATGAT TTGGATGACA CCAAGTATCA	2950
AGATCCCAAA TACTGGTCAA ATCCGTTAGA GACAAAGTAA GTGATTCTT	3000
GATTCCAACG GTATCCTTCG TCCTGATGCA TTATCAGTCT TTTTGTITTC	3050
GGTCTTGTG GATATGGTTT TCAGCTCAA GCTTACAAAG CTGTTCTGA	3100
GCCTTCTCA AAAAGGCTTG CTCAGTAATA TTGAGGTGCT AAAGTTGATA	3150
CATGTGACTC TTGCTTATAA ATCCCTCCGTT TGGTTGTTC TGCTTTTCA	3200
GATTACCGAA TGCTCCTGAG ATGGAAATCT ACTCATTATA CGGAGTGGGG	3250
ATACCAACGG AACGAGCATA CGTATACAAG CTTAACCAAGT CTCCCGACAG	3300
TTGCATCCCC TTTCAGATAT TCACTCTGC TCACCGAGGAG GACGAAGATA	3350
GCTGTCTGAA AGCAGGAGTT TACAATGTGG ATGGGGATGA AACAGTACCC	3400
GTCCTAAGTG CCCGGTACAT GTGTGCATAA GCGTGGCGTG GCAAGACAAG	3450
ATTCAACCCCT TCCGGAATCA ACACCTTATAT AAGAGAATAC AATCACTCTC	3500
CGCCGGCTAA CCTGTTGGAA GGGCGCGGGGCGCAGAGTGG TGCCCATGTT	3550
GATATCATGG GAAACTTTGC TTTGATCGAA GATATCATGA CGGTTGCCGC	3600
CGGAGCTAAC GGGCTGTATA TAGGACATGA CCAGGTCCAC TCTGGCATAT	3650
TGGAATGGTC GGAGCGTATT GACCTGAAGC TGTGA	3685

<210> 5
 <211> 402
 <212> cDNA
 <213> Arabidopsis thaliana
 <220>
 <221> CDS
 <222> (120) .. (402)
 <400> 5

AGAACAGCTCTTGTCTCT	CTCGACTGATCTAACAAATCC	CTAATCTGTGTTCTAAATTG	60
CTGGACGAGATTTGACAAAG	TCCGTATAGCTAACCTGGT	TTAATTTCAGTGACAGAT	119
ATG CCC CTT ATT CAT CGG AAA AAG CCG ACG CAG AAA CCA TCG ACG CCG	Met Pro Leu Ile His Arg Lys Lys Pro Thr Glu Lys Pro Ser Thr Pro		167
CCA TCT GAA GAG GTG GTG CAC GAT GAG GAT TCG CAA AAG AAA CCA CAC	Pro Ser Glu Val Val His Asp Glu Asp Ser Gln Lys Lys Pro His		215
GAA TCT TCC AAA TCC CAC CAT AAG NAA TCG AAC GGA GGA GGG AAG TGG	Glu Ser Ser Lys Ser His His Lys ??? Ser Asn Gly Gly Gly Lys Trp		263
TCG TGC ATC GAT TCT TGT TGT TGG TTC ATT GGG TGT GTG TGT GTA ACC	Ser Cys Ile Asp Ser Cys Cys Trp Phe Ile Gly Cys Val Cys Val Thr		311
TGG TGG TTT CTT CTC TTC CTT TAC AAC GCA ATG CCT GCG AGC TTC CCT	Trp Trp Phe Leu Leu Phe Tyr Asn Ala Met Pro Ala Ser Phe Pro		359
CAG TAT GTA ACG GAG CCG AAT CAC GNG TCC TTT GCC TTA CCC G	Gln Tyr Val Thr Glu Pro Asn His ??? Ser Phe Ala Leu Pro		402

<210> 6
 <211> 643
 <212> cDNA
 <213> Zea mays
 <220>
 <221> CDS
 <222> (1)..(402)
 <400> 6

CGG GAG AAA ATA GCT GCT TTG AAG GGG GGT GTT TAC TTA GCC GAT GGT	48
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly	
1 5 10 15	
GAT GAA ACT GTT CCA GTT CTT AGT GCG GGC TAC ATG TGT GCG AAA GGA	96
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly	
20 25 30	
TGG CGT GGC AAA ACT CGT TTC AGC CCT GCC GGC AGC AAG ACT TAC GTG	144
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val	
35 40 45	
AGA GAA TAC AGC CAT TCG CCA CCC TCT ACT CTC CTG GAA GGC AGG GGC	192
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly	
50 55 60	
ACC CAG AGC GGT GCA CAT GTT GAT ATA ATG GGG AAC TTT GCT CTA ATT	240
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile	
65 70 75 80	
GAG GAC GTC ATC AGA ATA GCT GCT GGG GCA ACC GGT GAG GAA ATT GGT	288
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly	
85 90 95	
GGC GAT CAG GTT TAT TCA GAT ATA TTC AAG TGG TCA GAG AAA ATC AAA	336
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys	
100 105 110	
TTG AAA TTG TAA CCTATGGGAA GTTAAAGAAG TGCCGACCCG TTTATTCGTTCC	391
Leu Lys Leu	
115	
AAAGTGTCT GCCTGAGTGC AACTCTGGAT TTTGCTTAAA TATTGTAATT TTTCACCGC	449
TTCATTGTC CCTTTGTCAA ATTTACATTT GACAGGACGC CAATGCGATA CGATGTTG	507
TACCGCTATT TTCAGCATTG TATATTAAC TGTACAGGTG TAAGTTGCAT TTGCCAGC	565
TGAAATTGTG TAGTCGTTTT CTTTACGATT TAATANCAAG TGGCGGAGCA GTGCCCCA	623
AGCNAAAAAAA AAAAAAAAAAA	643

<210> 7
 <211> 115
 <212> PRT
 <213> Zea mays
 <400> 7

Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
 1 5 10 15

Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly
 20 25 30

Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val
 35 40 45

Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly
 50 55 60

Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile
 65 70 75 80

Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly
 85 90 95

Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys
 100 105 110

Leu Lys Leu
 115

<210> 8
<211> 616
<212> cDNA
<213> Neurospora crassa
<400> 8

GGTGGCGAAG ACCANGGCAG AAGTTGGAGG CTAACGAGAA TGACNCTCGG	50
AGATGGATCT ACCCTCTAGA GACACCACTA CCNTTGACCC CAGCCTCAAG	100
GTNTACNGTT TNTATGGGTA GGAAGCCGAC GGAGCGAGCC TACATCTATC	150
TGGCGCCCGA TCCCAGGACG ACAACGCATC TTTAGATGAC GATCGATAACG	200
ACTTTGACTN AGGGGCACAT TGACCACGGT GTGATTTGG GCGAAGGCAG	250
TGGCACAGTG AACCTTATGA GTTGGGGTA CCTGTGCAAT AAGGGGTGGA	300
AAATGAAAGAG ATACAATCCT GCGGGCTCAA AAATAACCGT GGTGGAGATG	350
CCGCATGAAAC CAGAACGGTT CAATCCGAGA GGAGGGCCGA ATACGGCGGA	400
CTTAATATG TAGAAAAGGT TGAAATTAT GAAGAGTAAT TAAATAACGGC	450
ACATAGGTAA CTCAATAGTA TGACTAATTA AAAAAAAATT TTTTTCTAA	500
Aaaaaaaa aaaaaaa	516

DRAFT DRAFT DRAFT

<210> 9
 <211> 1562
 <212> genomic DNA
 <213> Arabidopsis thaliana
 <400> 9

ATGAAAAAAA	TATCTTCACA	TTATTGGTA	GTCATAGCGA	TACTCGTTGT,	50
GGTGACGATG	ACCTCGATGT	GTCAAGCTGT	GGGTAGCAAC	GTGTACCCTT	100
TGATTCTGGT	TCCAGGAAAC	GGAGGTAACC	AGCTAGAGGT	ACGGCTGGAC	150
AGAGAATACA	AGCCAAGTAG	TGTCTGGTGT	AGCAGCTGGT	TATATCCGAT	200
TCATAAGAACG	AGTGGTGGAT	GGTTTAGGCT	ATGGTTCGAT	GCAGCAGTGT	250
TATTGTCTCC	CTTCACCAGG	TGCTTCAGCG	ATCGAATGAT	GTTGTACTAT	300
GACCCCTGATT	TGGATGATTA	CCAAAATGCT	CCTGGTGTCC	AAACCCGGGT	350
TCCTCATTTTC	GGTCGACCA	AATCACTTCT	ATACCTCGAC	CCTCGTCTCC	400
GGTTAGTACT	TTCCAAGATA	TATCATTTCG	GGACATTGTC	ATAATGAACA	450
AAATAGACAT	AAATTTGGGG	GATTATTGTT	ATATCAATAT	CCATTTATAT	500
GCTAGTCGGT	AATGTGAGTG	TTATGTTAGT	ATAGTTAATG	TGAGTGTAT	550
GTGATTTTCC	ATTTAAATG	AAGCTAGAAA	GTTGTCGTTT	ATAATGTTG	600
CTATGTCATG	AGAATTATAA	GGACACTATG	TAAATGTAGC	TTAATAATAA	650
GGTTTGATTT	GCAGAGATGC	CACATCTTAC	ATGGAACATT	TGGTGAAGGC	700
TCTAGAGAAA	AAATGCGGGT	ATGTTAACGA	CCAAACCATC	CTAGGAGCTC	750
CATATGATTT	CAGGTACGGC	CTGGCTGCTT	CGGGCCACCC	GTCCCCGTGT	800
GCCTCACAGT	TCCTACAAGA	CCTCAAACAA	TTGGTGGAAA	AAACTAGCAG	850
CGAGAACGAA	GGAAAGCCAG	TGATACTCCT	CTCCCCTAGC	CTAGGAGGAC	900
TTTCGTCCCT	CCATTTCTC	AACCGTACCA	CCCCTTCATG	GCGCCGCAAG	950
TACATCAAAC	ACTTTGTTGC	ACTCGCTGCG	CCATGGGGTG	GGACGATCTC	1000
TCAGATGAAG	ACATTTGCTT	CTGGCAACAC	ACTCGGTGTC	CCTTTAGTTA	1050
ACCCCTTGCT	GGTCAGACGG	CATCAGAGGA	CCTCCGAGAG	TAACCAATGG	1100
CTACTTCCAT	CTACCAAAGT	GTTCACGAC	AGAACTAAAC	CGCTTGTGCGT	1150
AACTCCCCAG	GTAACTACA	CAGCTTACGA	GATGGATCGG	TTTTTGCGAG	1200
ACATTGGATT	CTCACAAAGGA	GTTGTGCCTT	ACAAGACAAG	AGTGTGCGCT	1250

TTAACAGAGG AGCTGATGAC TCCGGGAGTG CCAGTCACCT GCATATATGG	1300
GAGAGGAGTT GATACACCGG AGGTTTGAT GTATGGAAAA GGAGGATTCCG	1350
ATAAGCAACC AGAGATTAAG TATGGAGATG GAGATGGGAC GGTAAATTG	1400
GCGAGCTTAG CAGCTTGAA AGTCGATAGC TTGAACACCG TAGAGATTGA	1450
TGGAGTTTCG CATACATCTA TACTTAAAGA CGAGATCGCA CTTAAAGAGA	1500
TTATGAAGCA GATITCAATT ATTAATTATG AATTAGCCAA TGTTAATGCC	1550
GTCAATGANT GA	1562

<210> 10
<211> 3896
<212> genomic DNA
<213> Arabidopsis thaliana
<400> 10

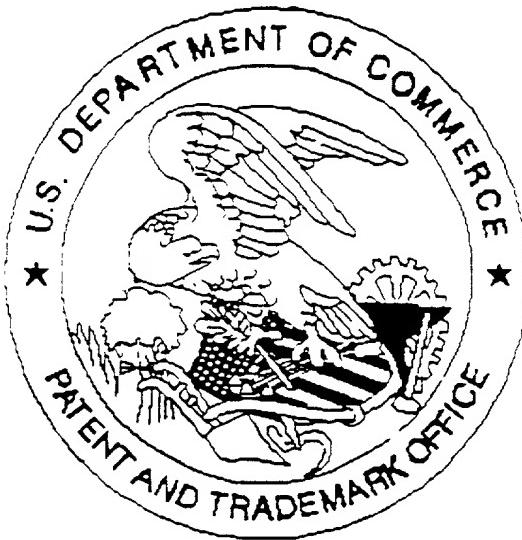
ATGGGAGCGA	ATTCGAAATC	AGTAACGGCT	TCCTTCACCG	TCATGCCGT	50
TTTTTCTTG	ATTTGGGCG	GCCGAACCTGC	GGTGGAGGAT	GAGACCGAGT	100
TTCACGGCGA	CTACTCGAAG	CTATGGGTA	TAATCATCGC	CCCGATTCCG	150
TCGACCGAGC	TACGAGCGTG	GTCGATCCTT	GAAGTCCCAT	ACACTCCGTT	200
CGACTTCAAT	CCGCTCGACC	TCGTATGGCT	AGACACCACT	AAGGTCCGTG	250
ATCTTCATT	CCTTCGCTCC	TTATTCTGTC	GGTCGAGTC	CTTGTGATG	300
AATTCCAAGC	GAAATAATAGC	AATGAAAGCAT	CTCTCGTCTC	TCTTATTGAT	350
TCGTTCAATT	GTCACAGTG	ACGCTTCTGA	ATCTGAGTTT	AGAGTCATAT	400
AAAACAGCTG	ACTCGGGCGAG	TCTTCCCCT	CGCTTTGGT	TCGCTAAATG	450
TAGCGCAATG	AATGTGTAAT	TAGTCTGCGC	TTTTTATTC	ACTAGATCTG	500
CAAGTTTTC	AGAGTGCCTA	ATAGTAGTIA	GAATAATGTTA	GGTCATTTA	550
CTTGTGCATT	GTGATTCTTT	GGTTGTTGTC	TTACTGATCG	ACGTGATGGA	600
TGGTTTACAG	CTTCCTTCG	CTGTCACCTG	CTGGTTAAG	TGTATGGTGC	650
TAGATCCTA	TARTCAAACA	GACCATCCC	AGTGTAAAGTC	ACGGCCTGAC	700
AGTGGCTTT	CAGGCATCAC	AGAATGGAT	CCAGGTACAC	TAACAGGTAG	750
TTTCGGATTT	TTCTTCTTT	TGAGTTTCT	TCATTGAT	ATCATCTTGT	800
TGTGATATA	TATGGCTAAG	TCATTAATT	TGGTCAATT	TCAGGTCC	850
TTTCTACTGT	CTGGAAAAGAG	TGGCTTAAGT	GGTGTGTTGA	CTTTGGTATA	900
GAAGCCAAATG	CAATTGTCG	TGTTCCATAC	GATTGGAGAT	TGTCACCAAC	950
CAATTGGAA	CGCGTGTGACC	TTTACTTTCA	CAAGCTCAAG	TTAGTCCTTA	1000
TCAGGCTAAT	GTCTTTATC	TTCTCTTTT	ATGTAAGATA	AGCTAAGAGC	1050
TCTGGTCGTC	TTCCCTTTTG	CAGGTTGACC	TTGAAACTG	CTTTAAAACT	1100
CCGTGGCGGC	CCTTCTATAG	TATTTGCCA	TTCAATGGGT	AATAATGTCT	1150
TCAGATACTT	TCTGGAATGG	CTGAGGCTAG	AAATTGCAAC	AAACACATTAT	1200
TTGAAGTGGC	TTGATCAGCA	TATCCATGCT	TATTTCGCTG	TGGTACCGG	1250
CCTACTATCC	TTAAGTTACC	ATTTTATTTT	TTCTCTAAT	GGGGGAGTTA	1300
TGTTGTGACT	TACTGGATTG	AGCTCGATAC	CTGATTTGTT	GTTGATTTAG	1350
GAGCTCCCT	TCTTGGITCT	GITGAGGCA	TCAAATCTAC	TCTCTCTGGT	1400
GTAACGTTTG	GCCTTCTGT	TTCTGAGGTG	ACCTCTGACT	TCTCTT	1450
TTTAAGTAGT	TGATATCAAC	CAGGCTTAT	AACTCACTGG	ATTTCTTT	1500
TGAAAGTATT	ACTTTGTTA	ATTGAACCTG	TGACGCGAT	ATGGTATCTG	1550
TAGATCTTGA	AGTGTAGTT	ATCAAAGAAC	ATATTGTGGG	TAGTACACCT	1600
GTCAGCGGCC	TTAGCTAATA	CAACCAAACC	ACATGTACAC	TGATTAGTT	1650
TTCAGATTAT	TATGGTAGAC	TTAAGTTGA	CAAGAAACTT	TGACTGAAAT	1700
CITTTATTT	TAATAGGCTA	TGATTGTTT	ATTGAAATCA	TGTGACATAT	1750
TGACATGCGC	TTCTCATGTT	TTTGTGTC	AAAGCTTCAG	GGAACTGCTC	1800
GTTGTTGTC	CAATTCTTT	GCGTCGTAT	TGIGGTTAT	GCCATTTC	1850
AAAGATTGCA	ACGGTATAAA	CACATTCTGG	ACGCATTTT	CTGGGGGTGC	1900
TGCNAAGAAA	GATAAGCGCG	TATACCACTG	TGATGAAGAG	GAATATCAAT	1950
CAAAATATTC	TGGCTGGCCG	ACAAATATTA	TAAACATTGA	AATTCCCTCC	2000
ACTAGCGGTT	AGACTCTGTA	TATGCAACTG	TAACACTAAC	AAAAGTTCA	2050
CCAAGAATGT	TCACTCTCAT	ATTTCTTCC	TTGATGTTG	ATCCATCAGT	2100
TACAGAAACA	GCTCTAGTCA	ACATGACCG	CTGGAATGT	GGCCTTCCC	2150
CCCTTTGTC	TTTCACAGCC	CGTGAACCTAG	CAGATGGAC	TCTTCAAA	2200
GCAATAGAAG	ACTATGACCC	AGATAGCAAG	ACGATGTTAC	ACCAAGTTAA	2250
GAAGTACGTA	CCTTCTTTG	TGATAAGAAA	TATTGCTCAT	CGATCATCAC	2300
TTCGCTGGCTT	CTTGTACGTC	AAATTGTTT	GTGTTAATCT	CTATATCAAT	2350
TGTTCAATATG	CTTGTCTTT	CTTACTATAA	GAAACAAAGTA	TAATCAGAAA	2400
CCTTATTATT	GATTATCAGT	TCTCTCTTA	TATTATGGAA	TGTCTTTTC	2450
GTTTACAGTT	ATGATGCAA	AAGGGGTAT	TTAGTTGAT	TGATTCTCTC	2500
ATTCTCTAGT	TTGTTTGAC	TAATAGCGTC	ATTTTGT	TTCTAGCAA	2550
TCTTTGTGAA	TTATATATA	CATGCTAACT	ATACTTTCA	GGTTGATCA	2600
TGATGACCC	GTTTTAATC	CTCTGACTCC	TTGGGAGAGA	CCACCTATAA	2650
AAAATGTATT	TTGCATATAT	GGTGCTCATC	TAAGACAGA	GGTATGATGC	2700
ATTCTCAATA	TCACATTATG	CGTGTACTTT	GTTATTATAT	TCCCCATTG	2750

GTTTGCATA TCTTTTGAA TTATGATTA TCTTCTCCCT TGCATCTTAT	2800
CCTATTAGC GTTAAAGGT A CTAAATGTAT GAAGCTGTCT GTCATAGGT	2850
GGTTATTACT TTGCCCAAG TGGCAAACCT TATCCTGATA ATTGGATCAT	2900
CACGGATATC ATTTATGAAA CTGAAGGTT C CTCGTGTCA AGGTAACTTT	2950
CCGCAATGGC ACAAGTAAAA CAGGAAGGCA AAGTCITCTG TATCAGTCTA	3000
GTGGCATGTT ATCTCAGTTC CATAAGCAAA TTATTAAACA ACTAAAATT	3050
AAGTACTTTT TTAICATTCC TTTTGAGCTT AGTGGATGAT CAGTGGCTTA	3100
AAGTGGGAAG AGGTGTTGCA TGAAACATGA CACTTGATC AAAGATAACT	3150
ACCAAAACAA AACTAACCCA TTTCTGAATT TCATAATTATT AGGAGTAGTC	3200
GTGCTTTAA AAAATTGTT TTAAGAAACC GAAAAAACTAG TTCATATCTT	3250
GATTGTCAA TATCTGCAGG TCTGGAACCTG TGTTGATGG AACGCTGGA	3300
CCTATAACTG GGGATGAGAC GGTAAGCTCA GAAGTTGGTT TTGAATTAT	3350
CTTCCTGCAA ACTACTGAAAG ACTAAGATAA TACTTGCTTC TGGAACACTG	3400
CTTGCTATGT TCTCTAGTAC ACTGCAATAT TGACTCTCCG CTACTTTAT	3450
TGATTATGAA ATTGATCTCT TATAGGTACC CTATCATTCA CTCTCTTGGT	3500
GCAAGAATTG GCTCGGACCT AAAGTTAAC A TAAACAATGGC TCCCCAGGTA	3550
CTCCTTTTA GTTCCCTCCTC TTATATAGAT CAAACTTTA GTGTACTTTT	3600
CTGGTTATGT GTTGATTAC CTCCATTG TTCTTTCTAA AAATCATATA	3650
TCTCTGTACT CCTCAAGAAC TTGTATTAAT CTAAACGAGA TTCTCATTGG	3700
GAATAATAAA CAACAGCCAG AACACGGATGG AAGCGACGTA CATGTGGAC	3750
TAAATGTTGA TCATGAGCAT GGGTCAGACA TCATAGCTAA CATGACAAAAA	3800
GCACCAAGGG TTAAGTACAT AACCTTTAT GAAGACTCTG AGAGCATICC	3850
GGGGAAAGAGA ACCCGAGTCT CGGAGCTTGA TAAAAGTGGG TATTAA	3896

<210> 11b
<211> 709
<212> cDNA
<213> tomato
<400> 11

CCTGGGGCCAA AAGTGAACAT AACAAAGGACA CCACAGTCAG ACCATGATGT	50
TCAGATGTAC AAGTGCATCT AAATATAGAG CATCAACATG GTGAAGATAT	100
CATTCCCCAT ATGACAAAAGT TACCTACAAT GAAGTACATA ACCTATTATG	150
AGGATTCTCA AAGTTTCCA GGGACAAGAA CAGCAGTTG GGAGCTTGAT	200
AAAGCAALTC ACAGGAACAT TGTCAGAICT CCACCTTGA TGCAGGGAGCT	250
GTGGCTTGAG ATGTGGCATG ATATTCACTCC TGATAAAAAG TCCAAGTTG	300
TTACAAAAGG TGGTGTCTGA TCCTCACTAT TTTCTTCTAT AAATGTTGA	350
GTTTGTATTG ACATTGTAAG TATTGCAACA AAAAGCAAAG CGTGGGCCTC	400
TGAGGGATCA GGACTGCTAT TGGGATTACG GGAAAGCTCG ATGTGCATGG	450
GCTGAACATT GTGAATACAG GTTAGAATAT TCAAATTATA TTTTGCAAAA	500
TATTCTCTTT TTGTGTATTT AGGCCACCTT TCCCCGGTCA CAACGATGCA	550
GATATGTATT CGGGGATGTT CACCTGGGAC AGAGTTGCAG ATTGAAGAGT	600
TCTACATCTC ACATCCTGTC ACACATGTG TGATATTTAA GAAACTTTGT	650
TTGGCGGAAC ACAAGTTG CACAAACATT TGAAGAAGAA ACCGAAATGA	700
TTCAGAGAG	709

United States Patent & Trademark Office
Office of Initial Patent Examination -- Scanning Division



SEARCHED INDEXED
SERIALIZED FILED

Application deficiencies were found during scanning:

Page(s) _____ of _____ were not present
for scanning. (Document title)

Page(s) _____ of _____ were not present
for scanning. (Document title)

Scanned copy is best available.

The first six pages of the sequence listing are the drawings.